

STIC-Biotech/ChemLib

73796

From: Chan, Christina
Sent: Tuesday, August 20, 2002 3:19 PM
To: Wilder, Cynthia; STIC-Biotech/ChemLib
Subject: RE: Rush sequence search 09/835,082

12 E11

Please rush. Thanks Chris

-----Original Message-----

From: Wilder, Cynthia
Sent: Tuesday, August 20, 2002 12:14 PM
To: Chan, Christina
Subject: Rush sequence search 09/835,082

Ms. Chan,

I am requesting a Rush sequence search for 09/835,082 due to a low docket. Please forward your approval to STIC.

Please search the nucleic acid encoded by SEQ ID NO: 1 and SEQ ID NO: 3.

Thank you!

Cynthia B. Wilder, Ph.D.

Art Unit 1637

703/305-1680

Office No. 12 F11

Mailbox No. 10 F12

Edward Hart
Technical Info. Specialist
STIC/Biotech
CMI 6B02 Tel: 305-9203

Searcher: _____
Phone: _____
Location: _____
Date Picked Up: 8/20/02
Date Completed: 8/21/02
Searcher Prep/Review: _____
Clerical: _____
Online time: _____

TYPE OF SEARCH:

NA Sequences: 2
AA Sequences: _____
Structures: _____
Bibliographic: _____
Litigation: _____
Full text: _____
Patent Family: _____
Other: _____

VENDOR/COST (where applic.)

STN: _____
DIALOG: _____
Questel/Orbit: _____
DRLink: _____
Lexis/Nexis: _____
Sequence Sys.: 05
WWW/Internet: _____
Other (specify): _____

[The body of the page contains extremely faint, illegible text, likely bleed-through from the reverse side of the paper. The text is too light to be transcribed accurately.]

•

Db 601 catatgattcttctgttgggaaagaaatgatgctatgaaattggtggtgtatt 660
 QY 2211 tgagaagataatcatgtcttctgcaaatgagctgaattgataaaacccaaatata 2270
 Db 661 tgagaagataatcatgtcttctgcaaatgagctgaattgataaaacccaaatata 720
 QY 2271 gctatgaagtctggtggcaagtttacttttttttctgagtttctactaac 2320
 Db 721 gctatgaagtctggtggcaagtttacttttttttctgagtttctactaac 770
 RESULT 11
 AA245116
 ID AA245116 standard; DNA; 930 BP.
 AC AA245116;
 XX 28-FEB-2030 (first entry)
 DE UDP-glucuronosyltransferase 1 (UGT1) exon 1H nucleotide sequence.
 KW Uridine diphosphate-glucuronosyltransferase 1; UGT1; polymorphism; probe;
 KW glucuronic acid; Crigler-Najjar syndrome; Gilbert syndrome; jaundice;
 KW unconjugated hyperbilirubinemia; drug metabolism; transgenic animal;
 KW pharmacogenetic screening; diagnosis; ss.
 XX Homo sapiens.
 XX W09957322-A2.
 XX 11-NOV-1939.
 XX 04-MAY-1999; 99MO-US09702.
 XX 07-MAY-1998; 98US-0084807.
 XX (AXYS-) AXYS PHARM INC.
 XX Penny L, Galvin M;
 XX WPI: 2000-052981/04.
 XX P-PSDB; AAY37098.
 XX New nucleic acid representing polymorphisms in the human uridine
 XX diphosphate glucuronosyltransferase gene, used for diagnosis and
 XX evaluation of drug metabolism -
 XX Examples: Page 39-41; 63pp; English.
 XX Nucleotide sequences AA245110-245118 are exons 1A-1J of the human
 XX uridine diphosphate-glucuronosyltransferase 1 (UGT1) gene. The UGTs are a
 XX family of enzymes that catalyse the glucuronic acid conjugation of a wide
 XX range of endogenous and exogenous substrates including phenols,
 XX alcohols, amines and fatty acids. Many of the reactions catalysed by UGTs
 XX result in toxic substances being converted to compounds which are more
 XX water soluble and are excreted. The invention relates to and identifies
 XX UGT1 polymorphisms (AA245004-245041). The polymorphism sequences are
 XX useful as probes for detecting UGT1 locus polymorphisms, indicative of
 XX altered UGT1 expression or activity. These polymorphisms are associated
 XX with Crigler-Najjar and Gilbert syndromes (unconjugated
 XX hyperbilirubinemia) and drug metabolism. The genotyping of the UGT1 gene
 XX is used to predict the rate of metabolism of UGT1 substrates, possible
 XX drug-drug interactions and adverse side effects (i.e. to optimize drug
 XX dosage), and to screen for diseases caused by exposure to toxins and to
 XX study the effects of polymorphisms on enzymatic activity. The UGT1
 XX sequences, including polymorphisms, can also be used to produce the
 XX corresponding protein (or its fragments) or to generate transgenic
 XX animals or modified cells e.g. for pharmacogenetic screening.
 XX Sequence 930 BP; 228 A; 205 C; 212 G; 285 T; 0 other;

Query Match 33.0%; Score 766.6; DB 21; Length 930;
 Best Local Similarity 91.6%; Pred. No. 2.1e-207;
 Matches 811; Conservative 0; Mismatches 74; Indels 0; Gaps 0;
 QY 19 atggcttgcaacaggggtgacacagcccccttccctctatgtgtgtctgtctgacctgt 78
 Db 1 atggcttgcaacaggggtgacacagcccccttccctctatgtgtgtctgtctgacctgt 60
 QY 79 ggccttccagagcaggggaagctactgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 138
 Db 61 ggccttccagagcaggggaagctactgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 120
 QY 139 atgaggtcggt 198
 Db 121 atgaggtcggt 180
 QY 199 gaggtgagttggcaactgggaagatcaactgaattgcacagtgaaagacttattcaactca 258
 Db 181 gaggtgagttggcaactgggaagatcaactgaattgcacagtgaaagacttattcaactca 240
 QY 259 tataacctgagagatctgacacggaggttcaagcttttgcacatgctcaatgaaagca 318
 Db 241 tataacctgagagatctgacacggaggttcaagcttttgcacatgctcaatgaaagca 300
 QY 319 caagtacgaagtatact 378
 Db 301 caagtacgaagtatact 360
 QY 379 ttttcaaatgacagaggtttgtttaaagacaaaataatgtagaatactttaaagagaggt 438
 Db 361 ttttcaaatgacagaggtttgtttaaagacaaaataatgtagaatactttaaagagaggt 420
 QY 439 tcttttgatcgagttct 498
 Db 421 tcttttgatcgagttct 480
 QY 499 tttccctccctccctggt 558
 Db 481 tttccctccctccctggt 540
 QY 559 gcacagtgccctgct 618
 Db 541 gcacagtgccctgct 600
 QY 619 atgactttcaagagagagtagacggaacacacacacacacacacacacacacacacacac 678
 Db 601 atgactttcaagagagtagacggaacacacacacacacacacacacacacacacacac 660
 QY 679 caccgttttttcaaaaatgccttagaataagcctgtgaaattctccaaacacacacacac 738
 Db 661 caccgttttttcaaaaatgccttagaataagcctgtgaaattctccaaacacacacacac 720
 QY 739 gagtatgctctacagcagac 798
 Db 721 gagtatgctctacagcagac 780
 QY 799 tatcccaaacccgtgtagcccaacacacacacacacacacacacacacacacacacacac 858
 Db 781 tatcccaaacccgtgtagcccaacacacacacacacacacacacacacacacacacacac 840
 QY 859 aagcgttgcctatggaattggaagcctacattatgcttctgtgga 903
 Db 841 aagcgttgcctatggaattggaagcctacattatgcttctgtgga 885
 RESULT 12
 AA245115
 ID AA245115 standard; DNA; 951 BP.
 AC AA245115;
 XX 28-FEB-2000 (first entry)


```

Db 541 gacgtcacagaggctgcattctagattactttcttacttgaaacatggcctgttg 600
Qy 1922 ggaagtgcggattcaaaaggtggtccacacgctgccctactgcacaaatggcagtttaac 1981
Db 601 ggaagtgcggattcaaaaggtggtccacacgctgccctactgcacaaatggcagtttaac 660
Qy 1982 ttatctttggtcttcagatggttgcaattgacattacccaataatggtcagtcctc 2041
Db 661 ttatctttggtcttcagatggttgcaattgacattacccaataatggtcagtcctc 720
Qy 2042 atctctgtctcttcattagtgccacactgtgtgttttaagaagggaagctttgtacct 2101
Db 721 atctctgtctcttcattagtgccacactgtgtgttttaagaagggaagctttgtacct 780
Qy 2102 ttgaagttagtggaatgaatgaatgcttggagtgacactgagacagcgcattgattc 2161
Db 781 ttgaagttagtggaatgaatgaatgcttggagtgacactgagacagcgcattgattc 840
Qy 2162 ttgctttgggaaaaaagaatgatctatgaatgggtgggtgtatttgagaagataa 2221
Db 841 ttgctttgggaaaaaagaatgatctatgaatgggtgggtgtatttgagaagataa 900
Qy 2222 tcattcttatgtcaaatgagctgaatttgataaaacccaataacacgctatgaagtg 2281
Db 901 tcattcttatgtcaaatgagctgaatttgataaaacccaataacacgctatgaagtg 960
Qy 2282 ctggggcaagtttactttttttctgtatgtttctctacaact 2320
Db 961 ctggggcaagtttactttttttctgtatgtttctctacaact 999

```

RESULT 8

```

AAH51404
ID AAH51404 standard; DNA; 1001 BP.
XX
AC AAH51404;
XX
DT 29-AUG-2001 (first entry)
XX
DE Human UGT1A7 related DNA containing a biallelic polymorphism SEQ ID 295.
XX
KW Human; biallelic marker; single nucleotide polymorphism; SNP; MGSTII;
KW microsomal glutathione S-transferase II; malate decarboxylase enzyme;
KW DME1; ME1; cytochrome P450; glutathione reductase; GSHR; GSHS; GGT5;
KW flavin-containing monooxygenase; FMO; gamma-glutamyltransferase 5;
KW dipeptidase; DP; glucose 6-phosphate dehydrogenase; G6PDH; haplotype;
KW phosphogluconate dehydrogenase; PGDH; drug metabolism; phenotype;
KW uridine diphosphate glucuronosyl transferase; UGT2; asthma; hepatotoxicity;
KW zileuton; ds.
XX
OS Homo sapiens.
XX
PN W0200058508-A2.
XX
PD 05-OCT-2000.
XX
PF 24-MAR-2000; 2000WO-IB00403.
XX
PR 25-MAR-1999; 99US-0126269.
PR 30-APR-1999; 99US-0131961.
XX
PA (GEST ) GENSET.
XX
PI Blumenfeld M, Bougueleret L, Chumakov I, Cohen-Akenine A;
XX
DR WPI; 2000-638353/61.
XX
PR Polynucleotides comprising sequences from malate decarboxylase
PR enzyme-related biallelic markers used for genotyping -
XX
PS Claim 13; Page 487-488; 673pp; English.
XX

```

```

CC Sequences AAH51110-AAH51593 represent human DNA fragments which contain
CC biallelic markers. The sequences are related to various human genes
CC including microsomal glutathione S-transferase II (MGSTII), malate
CC decarboxylase enzyme (DME1/ME1), cytochrome P450, glutathione
CC reductase/synthase (GSHR/GSHS), flavin-containing monooxygenases (FMO),
CC gamma-glutamyltransferase 5 (GGT5), dipeptidase (DP), glucose 6-phosphate
CC dehydrogenase (G6PDH), phosphogluconate dehydrogenase (PGDH), and uridine
CC diphosphate glucuronosyl transferases (UGT2). Each of these sequences
CC contains a biallelic marker/polymorphism, which is represented in the
CC sequence as a degenerate/undefined base. The genes to which the biallelic
CC marker containing sequences are related are involved in drug metabolism.
CC Sequences AAH51594 - AAH51598 represent the genomic sequence of the
CC MGSTII gene and four alternative MGSTII cDNA sequences. AAH52905-AAH52906
CC are MGSTII gene products. PCR primers AAH51599 and AAH51600 are used in
CC an example for the amplification of human genomic DNA fragments. The
CC inventory includes a method of genotyping comprising determining the
CC identity of a nucleotide at a DME- or MGSTII-related biallelic marker in
CC a biological sample. The method is used to determine the frequency in
CC population of an allele of a DME- or MGST-II related biallelic marker and
CC to select an individual for inclusion in a clinical trial of a drug
CC treatment. The method is also used to detect association between allele
CC and phenotype, and to detect association between haplotype and phenotype.
CC The polynucleotides are used, in hybridization assays, sequencing assays
CC or allele specific amplification assays. The method can be used to
CC determine whether an individual suffers or is at risk of developing
CC asthma or is at risk of developing hepatotoxicity on treatment with
CC zileuton.
XX
SQ Sequence 1001 BP; 293 A; 171 C; 210 G; 326 T; 1 other;

```

Query Match 37.5%; Score 870.6; DB 21; Length 1001;

Best Local Similarity 99.9%; Pred. No. 5.6e-237;

Matches 870; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1450 ctcacgtgtaccagtaaccattctcttgacgtgatgttctcttggcgtgctgctg 1509

Db 1 ctcacgtgtaccagtaaccattctcttgacgtgatgttctcttggcgtgctgctg 60

Qy 1510 acagtggtctcatcaccctttaaattgtgtgtcttatggtaccggaaatgcttgggaaa 1569

Db 61 acagtggtctcatcaccctttaaattgtgtgtcttatggtaccggaaatgcttgggaaa 120

Qy 1570 aaagggcagtgtaagaagcccaaaatccaaagccattgagaagtggtgggaaataa 1629

Db 121 aaagggcagtgtaagaagcccaaaatccaaagccattgagaagtggtgggaaataa 180

Qy 1630 ggtaaaatttgaaacattcccttagctatcttccaaacttgaaacacagacagtgtaaa 1689

Db 181 ggtaaaatttgaaacattcccttagctatcttccaaacttgaaacacagacagtgtaaa 240

Qy 1690 ttcatatttattcttattaaaggaaatactttgtcataataatcagccagagtgcttta 1749

Db 241 ttcatatttattcttattaaaggaaatactttgtcataataatcagccagagtgcttta 300

Qy 1750 aaaaattctcttaaaaaataatagactcgtagctagtaagaatatttgaaatgta 1809

Db 301 aaaaattctcttaaaaaataatagactcgtagctagtaagaatatttgaaatgta 360

Qy 1810 tcgtgccctcccggtgtctttgatcagatgacatgtgccatttttcagagacgtgca 1869

Db 361 tcgtgccctcccggtgtctttgatcagatgacatgtgccatttttcagagacgtgca 420

Qy 1870 gacagctggcattctagattactttcttactctgaaacatggcctgtttggagtcg 1929

Db 421 gacagctggcattctagattactttcttactctgaaacatggcctgtttggagtcg 480

Qy 1930 ggattcaagtggtgccaccgctgccctactgcaaatggcagtttttaattcttcttt 1989

Db 481 ggattcaagtggtgccaccgctgccctactgcaaatggcagtttttaattcttcttt 540

Qy 1990 tggctctgcagatggttgcattgaccttaaccataatggtcagtcctctctctgt 2049


```

XX FH Key Location/Qualifiers
FT variation replace (2510,C)
FT /tag= a
FT /standard_name= "Single nucleotide polymorphism"
FT CDS 2616..14887
FT /tag= b
FT /product= "UDP glycosyltransferase 1"
FT /note= "this CDS is specifically claimed in claim
FT 26"
FT variation replace (2756,T)
FT /tag= c
FT /standard_name= "Single nucleotide polymorphism"
FT variation replace (2826,A)
FT /tag= d
FT /standard_name= "Single nucleotide polymorphism"
FT variation replace (3155,G)
FT /tag= e
FT /standard_name= "Single nucleotide polymorphism"
FT variation replace (9508,C)
FT /tag= f
FT /standard_name= "Single nucleotide polymorphism"
FT variation replace (9511,T)
FT /tag= g
FT /standard_name= "Single nucleotide polymorphism"
FT variation replace (10091,T)
FT /tag= h
FT /standard_name= "Single nucleotide polymorphism"
FT variation replace (10094,T)
FT /tag= i
FT /standard_name= "Single nucleotide polymorphism"
FT variation replace (10095,C)
FT /tag= j
FT /standard_name= "Single nucleotide polymorphism"
FT variation replace (10140,C)
FT /tag= k
FT /standard_name= "Single nucleotide polymorphism"
FT variation replace (14423,T)
FT /tag= l
FT /standard_name= "Single nucleotide polymorphism"
FT variation replace (14713,T)
FT /tag= m
FT /standard_name= "Single nucleotide polymorphism"
FT variation replace (14776,T)
FT /tag= n
FT /standard_name= "Single nucleotide polymorphism"
FT variation replace (14971,C)
FT /tag= o
FT /standard_name= "Single nucleotide polymorphism"
XX WO200179230-A2.
XX PN
XX PD
XX PP
XX PR
XX PR
XX PA (GENA-) GENAISSANCE PHARM INC.
XX PI Chew A, Choi JY, Koshy B, Rounds E;
XX PI WPI: 2002-075063/10.
XX DR P-PSDB; AAU10940.
XX DR
XX PT Genotyping a human UDP glycosyltransferase 1 gene of an individual for
XX PT determining the haplotype of an individual, involves determining the
XX PT identity of a nucleotide pair at specific polymorphic sites for two
XX PT copies of the gene
XX PS Claim 21; Fig 1; 81pp; English.
XX PS
XX XX The invention relates to genotyping a human UDP glycosyltransferase

```

```

CC (UGT1A1) gene of an individual, involving determining for the two copies
CC of the UGT1A1 gene present in the individual, the identity of the
CC nucleotide pair at one or more polymorphic sites. The new method is
CC useful for determining whether an individual has a haplotype or haplotype
CC pairs, given in the specification. It is useful for improving the
CC efficacy and reliability of several steps in the discovery and
CC development of drugs for treating diseases associated with UGT1A1
CC activity, e.g., Gilbert's syndrome and Crigler-Najjar syndrome, to
CC validate UGT1A1 as a candidate agent for treating a specific condition or
CC disease predicted to be associated with UGT1A1 activity, and in the
CC design of clinical trials of candidate drugs for treating a specific
CC condition or disease predicted to be associated with UGT1A1 activity. The
CC method is useful to screen for compounds targeting UGT1A1 activity. A nucleic
CC specific condition or disease associated with UGT1A1 activity. A nucleic
CC acid (I) comprising a polymorphic variant of a reference sequence for
CC the UGT1A1 gene or cDNA (II) or its fragment is useful in studying the
CC expression and function of UGT1A1, and in expressing UGT1A1 protein for
CC use in screening for candidate drugs to treat diseases related to UGT1A1
CC activity. (I) or (II) is useful for therapeutic purposes. (II) or a
CC recombinant organism comprising (II) is useful for studying
CC expression of the UGT1A1 isogenes in vivo, for in vivo screening and
CC testing of drugs targeted against UGT1A1 protein, and for testing the
CC efficacy of therapeutic agents and compounds for Gilbert's syndrome and
CC Crigler-Najjar syndrome, in a biological system. The present sequence
CC represents the coding sequence of UDP glycosyltransferase 1.
XX
SQ Sequence 18887 BP; 5079 A; 4037 C; 4131 G; 5640 T; 0 other;

Query Match 43.3%; Score 1004.4; DB 24; Length 18887;
Best Local Similarity 98.4%; Pred. No. 2.8e-274;
Matches 1014; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

QY 1291 aaagcagtcatacatgacaaaagtacaaagagagacatcatcgctctccagcttcac 1350
DB 14567 aacgaactgtcttggtagtttagtaacaaagagacatcatcgctctccagcttcac 14626
QY 1351 aaggaaccccggtggagccgctggacctggccgtgtcttgggtggagtgtgatgagg 1410
DB 14627 aaggaccgcccgggtggagccgctggacctggccgtgtcttgggtggagtgtgatgagg 14686
QY 1411 cacaaggcgccgcacacacctgcccgcgcagccacacacctggttaccagttaccat 1470
DB 14687 cacaaggcgccgcacacacctgcccgcgcagccacacacctggttaccagttaccat 14746
QY 1471 tccttgagcgtgattgttctcttcttgcctctgacctctgctgacagtgccttcacacctt 1530
DB 14747 tccttgagcgtgattgttctcttcttgcctctgacctctgctgacagtgccttcacacctt 14806
QY 1531 aaatgttgtcttatgctgaccgaaatgcttggggaaaaaaggcgaggttaagaaagcc 1590
DB 14807 aaatgttgtcttatgctgaccgaaatgcttggggaaaaaaggcgaggttaagaaagcc 14866
QY 1591 cacaatccaagaccattgagaagtgggtgggaaataaaggttaaaatttgaaccattcc 1650
DB 14867 cacaatccaagaccattgagaagtgggtgggaaataaaggttaaaatttgaaccattcc 14926
QY 1651 ctagtcatctccaactgaaacagaatcagtggttaaatcattttattttatttaaggg 1710
DB 14927 ctagtcatctccaactgaaacagaatcagtggttaaatcattttattttatttaaggg 14986
QY 1711 aaatactttgcataaaataataacagccagagctgtttaaataattctttaaataaaaa 1770
DB 14987 aaatactttgcataaaataataacagccagagctgtttaaataattctttaaataaaaa 15046
QY 1771 taatagactcgtagtcagtaaaagatatattgaatatgtatcgtgcccccctcggtgtctt 1830
DB 15047 taatagactcgtagtcagtaaaagatatattgaatatgtatcgtgcccccctcggtgtctt 15106
QY 1831 tgatcagatgacatgtgccatttttcagagcagtcgacagcagctgcttcatttagatt 1890
DB 15107 tgatcagatgacatgtgccatttttcagagcagtcgacagcagctgcttcatttagatt 15166

```


QY 858 aagccgctgctcttggaattgaaacctaaattgcttcttgagaaacatggaattgt 917
DB 881 ggaagcactactaggaattgaaacctaaattgcttcttgagaaacatggaattgt 940
QY 918 ggtttctcttgagatcaatgctcagaataatccagagaaagaaagctatgcaattgc 977
DB 941 ggtttctcttgagatcaatgctcagaataatccagagaaagaaagctatgcaattgc 1000
QY 978 tgatgtttgggcaaaatccctcagacagctcctgtggcgtgacactggaacccagaccatc 1037
DB 1001 tgatgtttgggcaaaatccctcagacagctcctgtggcgtgacactggaacccagaccatc 1060
QY 1038 gaactctgcaacaacagatcttgaatgctcaccacaaacagatctcgtctgtca 1097
DB 1061 gaactctgcaacaacagatcttgaatgctcaccacaaacagatctcgtctgtca 1120
QY 1098 ccgagacccctgcttaccacccatgctgttcccatggtgtttatgaaagcatatg 1157
DB 1121 ccgagacccctgcttaccacccatgctgttcccatggtgtttatgaaagcatatg 1180
QY 1158 caatgctgtccatggtgatgacctgtgtgtgtgacagatgacaatgcaaacg 1217
DB 1181 caatgctgtccatggtgatgacctgtgtgtgtgacagatgacaatgcaaacg 1240
QY 1218 catgagactaaggagctgagtgacctgaatgttcttgaaatgactctctgaagattt 1277
DB 1241 catgagactaaggagctgagtgacctgaatgttcttgaaatgactctctgaagattt 1300
QY 1278 agaaatgcttaaaagcagtcacatcaatacaaaagtacaaagaaacatcatgcct 1337
DB 1301 agaaatgcttaaaagcagtcacatcaatacaaaagtacaaagaaacatcatgcct 1360
QY 1338 ctccagcttcaaaagcagtcacatcaatacaaaagtacaaagaaacatcatgcct 1397
DB 1361 ctccagcttcaaaagcagtcacatcaatacaaaagtacaaagaaacatcatgcct 1420
QY 1398 gttgtgatgagcaagagcgccacacacctgcgcgcagccacagaccctcacctg 1457
DB 1421 gttgtgatgagcaagagcgccacacacctgcgcgcagccacagaccctcacctg 1480
QY 1458 gtaccagatccattctcttgagctgattggtttctcttcttggcctgctgacagtgcc 1517
DB 1481 gtaccagatccattctcttgagctgattggtttctcttcttggcctgctgacagtgcc 1540
QY 1518 ctctacacatttaaatgtgtcttatgctacogaaatgcttgggaaagaaagcg 1577
DB 1541 ctctacacatttaaatgtgtcttatgctacogaaatgcttgggaaagaaagcg 1600
QY 1578 agttaagaagcccaaatccaaagaccattgagaagtggtgggaaataaggtaaaat 1637
DB 1601 agttaagaagcccaaatccaaagaccattgagaagtggtgggaaataaggtaaaat 1660
QY 1638 ttgtaacattccctagctcattccaaacttgaacacagatcagtttaattcaattt 1697
DB 1661 ttgtaacattccctagctcattccaaacttgaacacagatcagtttaattcaattt 1720
QY 1698 attctattaaagaaatcttgcataaaatcaatcagcccgagtgctttaaataatc 1757
DB 1721 attctattaaagaaatcttgcataaaatcaatcagcccgagtgctttaaataatc 1779
QY 1758 tcttaataaaataatgactcagctagtcagtaagatatattgatatgtatgctgcc 1817
DB 1780 tcttaataaaataatgactcagctagtcagtaagatatattgatatgtatgctgcc 1839
QY 1818 cctccggtgctttgatcagatgacatgaccttttccattttcaaggagcgctgcagacgct 1877
DB 1840 cctccggtgctttgatcagatgacatgaccttttccattttcaaggagcgctgcagacgct 1899
QY 1878 ggcattcagattacttcttactctgaacatgacctgtttggagtgccgattcaaa 1937
DB 1900 ggcattcagattacttcttactctgaacatgacctgtttggagtgccgattcaaa 1959

QY 1938 agtgggtccacccgctccctactgcaaaatggcagtttaattatcttcttggctct 1997
DB 1960 agtgggtccacccgctccctactgcaaaatggcagtttaattatcttcttggctct 2019
QY 1998 gcagatggttgcaattgatcttcaaccaataatggtcagtcctcatctctgctcttc 2057
DB 2020 gcagatggttgcaattgatcttcaaccaataatggtcagtcctcatctctgctcttc 2079
QY 2058 ataggtgcccac-ttgtgtctttaaagaaggagctttgtacccttagagtgagtgga 2116
DB 2080 ataggtgcccac-ttgtgtctttaaagaaggagctttgtacccttagagtgagtgga 2139
QY 2117 aatgaatgaatggttgagtgcaactgagacagacatgatttctgtcttggggaaaa 2176
DB 2140 aatgaatgaatggttgagtgcaactgagacagacatgatttctgtcttggggaaaa 2199
QY 2177 agaatgagctatgaaattggtggtggtggtggttatgtgagaagataaatcattgcttatgca 2236
DB 2200 agaatgagctatgaaattggtggtggtggttatgtgagaagataaatcattgcttatgca 2259
QY 2237 aatggagctgaattgataaaaccccaaaacacagctatgagtgctgggcaagttaact 2296
DB 2260 aatggagctgaattgataaaaccccaaaacacagctatgagtgctgggcaagttaact 2319
QY 2297 tttttctgagtggttctctacaact 2320
DB 2320 tttttctgagtggttctctacaact 2343

RESULT 5

AAS91252

ID AAS91252 standard; cDNA; 1655 BP.

XX AC AAS91252;

XX XX 13-FEB-2002 (first entry)

XX DE DNA encoding novel human diagnostic protein #27056.

XX KW Human; chromosome mapping; gene mapping; gene therapy; forensic;

XX KW food supplement; medical imaging; diagnostic; genetic disorder; ss.

XX OS Homo sapiens.

XX PN WO200175067-A2.

XX PD 11-OCT-2001.

XX PF 30-MAR-2001; 2001WO-US08631.

XX PR 31-MAR-2000; 2000US-0540217.

XX PR 23-AUG-2000; 2000US-0649167.

XX PA (HYSE-) HYSEQ INC.

XX PI Drmanac RT, Liu C, Tang YT;

XX DR WPI; 2001-639362/73.

XX DR P-PSDB; ABG27065.

XX PT New isolated polynucleotide and encoded polypeptides, useful in

XX PT diagnostics, forensics, gene mapping, identification of mutations

XX PT responsible for genetic disorders or other traits and to assess

XX PT biodiversity

XX PS Claim 1; SEQ ID No 27056; 103pp; English.

XX CC The invention relates to isolated polynucleotide (I) and

XX CC polypeptide (II) sequences. (I) is useful as hybridisation probes,

XX CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome

XX CC and gene mapping, and in recombinant production of (II). The

XX CC polynucleotides are also used in diagnostics as expressed sequence tags

XX CC for identifying expressed genes. (I) is useful in gene therapy techniques

Db 1870 gtgcagacagctggcattctagattacttttcttactctgaaacatggcgtgtttggga 1929
 QY 1925 gtgcggattcaaaaggtgtccaccgcgtgcccactactgcaaatggcagttttaacttta 1984
 Db 1930 gtgcggattcaaaaggtgtccaccgcgtgcccactactgcaaatggcagttttaacttta 1989
 QY 1985 tcttcttgcttgcagatggttgcgaattgatcttcaacaaataatggcagtcctcatc 2044
 Db 1990 tcttcttgcttgcagatggttgcgaattgatcttcaacaaataatggcagtcctcatc 2049
 QY 2045 tctgctctgcttcaatagggtccacc-ttgtgtgttttaagaagggaagctttgtacctt 2103
 Db 2050 tctgctctgcttcaatagggtccacc-ttgtgtgttttaagaagggaagctttgtacctt 2109
 QY 2104 agag-gtagtggaatgaatgaatgcttggagtgcaactgagacacacatgatttctt 2163
 Db 2110 agag-gtagtggaatgaatgaatgcttggagtgcaactgagacacacatgatttctt 2169
 QY 2164 gctt-ggggaaaaagaatgctgctatgaaattgtgggtgtgtatttgagaagataatc 2223
 Db 2170 gctt-ggggaaaaagaatgctgctatgaaattgtgggtgtgtatttgagaagataatc 2229
 QY 2224 attgcttatgtcaaatgagctgaatttgataaaaccccaaaatacacatgaaagtgtc 2283
 Db 2230 attgcttatgtcaaatgagctgaatttgataaaaccccaaaatacacatgaaagtgtc 2289
 QY 2284 gggcnaagttacttttttctgattgtttctcacaact 2320
 Db 2290 gggcnaagttacttttttctgattgtttctcacaact 2326

RESULT 4
 AAQ27370
 ID AAQ27370 standard; cDNA; 2368 BP.
 XX AC AAQ27370;
 XX DT 27-JAN-1993 (first entry)
 XX DE HUG-Br2.
 XX KW Billrubin: UDP-glucuronosyltransferase; HUGBr1; HUGBr2;
 XX KW monoglucuronide; diglucuronide; ss.
 XX OS Homo sapiens.
 XX FH Key Location/Qualifiers
 XX FT CDS 30..801
 XX FT /*tag= a
 XX FT polyA_signal 2347..2352
 XX FT /*tag= b
 XX FT /*number= 1
 XX FT polyA_signal 2355..2360
 XX FT /*tag= c
 XX FT /*number= 2
 XX PN W09212987.A.
 XX PD 06-AUG-1992.
 XX XX 10-JAN-1992; 92NO-US00282.
 XX PF 10-JAN-1991; 91US-0639453.
 XX PR (USSH) US DEPT HEALTH & HUMAN SERVICE.
 XX PA Owens IS, Ritter JK;
 XX PI WPI; 1992-284593/34.
 XX DR P-PSDB; AAR26154.
 XX XX Isolated gene locus UGrl, DNA segments and diagnostic probes -
 XX PT for diagnosing Gilbert's disease and Crigler-Najjar syndrome

PT types I and II

XX Disclosure; Fig 9A-I; 99pp; English.

CC Two human liver bilirubin UDP-glucuronosyltransferase cDNAs have been isolated. They are referred to as HUGBr1 (AAQ27369) and HUGBr2 (AAQ27370) (Ritter, et al., J. Biol. Chem. 266:1043-1047 (1991)) and, upon expression individually in COS-1 cells, encode isoforms that catalyse the formation of the two bilirubin monoglucuronides and the diglucuronide.

CC The cDNAs contain identical 3' ends (1469 bp in length) to each other and to that of the human phenol transferase cDNA, HUGP1 (Harding et al., Proc. Natl. Acad. Sci. USA 85:8281 (1988)).

CC In contrast, they have unique 5' ends.

XX Sequence 2368 BP; 609 A; 532 C; 566 G; 661 T; 0 other;

Query Match 68.0%; Score 1576.8; DB 13; Length 2368;

Best Local Similarity 82.3%; Pred. No. 0;

Matches 1847; Conservative 0; Mismatches 392; Indels 5; Gaps 3;

QY 81 ctttgcgagcgagggaagctactgttagtgcccatggtgagcgccactggttaccat 140

Db 101 ctggcgctgaggtggaaaggtgtgtgtgcccactgatgagccctggtctcagcat 160

QY 141 gaggtcggtgtggagaactcattctcaggggcatgaggtgtgttagtcacagaga 200

Db 161 gcgggaggccttgcggagctccatgcagagggccaccaggcggtgctccaccacaag 220

QY 201 ggtgagttgcgaactgggaagatcactgaatgcacagtgcaagacttattcaactcata 260

Db 221 aggtgaatatgcacatcaagaagaattttccacctgacacgctctgctgtccatg 280

QY 261 taccctggaggtatctggaccgggagttcaaggcttttgcctatgctcaat--ggaaagc 317

Db 281 gaccagaaggaaatttgatcggttacgtgggtacactcaagggtctttgaaacaga 340

QY 318 acaagtcaggaagtattctctatttaagggttcatacaatgacatttttgacttatt 377

Db 341 acattctctgaagagatatcttagaagtaggcaattatgacaattatctttggccct 400

QY 378 ttttcaaatgcaggagttgttttaagacacaaaaattagtagaactttaaaggagag 437

Db 401 tcataggtgtgtgtgagctactgcataatgagccctgatcaggcacctgaatgctac 460

QY 438 tcttttgatgcagtttctcgatcccttttgataactgtgcttaattgtgccaaata 497

Db 461 tctcttgatgtgtgttttaacagaccccttcaacctctgcggggcgctgctaaagta 520

QY 498 tttctcctccctccgtgtgttcttgcgcagggggaatactttgcccactatcttgaagaagg 557

Db 521 cctgtcattctctgtgttttttggagggtacattccatgacttagactttaagg 580

QY 558 tgcagatgcttct 617

Db 581 cacacagttccaaatccct 640

QY 618 catgactttcaaggagagtagcaggaaccacatcatgcaacttgaggagaacatttatg 677

Db 641 catgacattctctcaagggttcaagaacatgctctaccctgtgcccctgtctctacattg 700

QY 678 ccacgcttttttcaaaaaatgccttagaataagctcttgaaattctccaaacacactgttac 737

Db 701 ccatacttttttgcct 760

QY 738 ggagtatgatctctacagccacacatcaatttggttgtgcgaacgactgttttttga 797

Db 761 tgtccaggacctattgagctctgcatctgtctgtgttttagaagtgactttgtgaagga 820

QY 798 ctatcccaaccccgtagtgcacacatgactctctctctctctctctctctctctctct 857

Db 821 ttaccctaggcccatcatgcccaatatggtcttctattgggggggcatcaactgttgcacaag 880

Qy 781 acggacttgggttttgactatcccaaacccggtgatgcccaacatgatcttcattggtgt 840
 Db 781 acggacttgggttttgactatcccaaacccggtgatgcccaacatgatcttcattggtgt 840
 Qy 841 atcaactgccatcaggaagcgttgcttatgaaatttgaagcctacattaatgctct 900
 Db 841 atcaactgccatcaggaagcgttgcttatgaaatttgaagcctacattaatgctct 900
 Qy 901 gagaaacatgaaattggttttctcttttggatcaatggtctcagaaattccagagaag 960
 Db 901 gagaaacatgaaattggttttctcttttggatcaatggtctcagaaattccagagaag 960
 Qy 961 aaagctatggcaattgctgatcttggtgcaaaatccctcagacagctcctgtgaggtac 1020
 Db 961 aaagctatggcaattgctgatcttggtgcaaaatccctcagacagctcctgtgaggtac 1020
 Qy 1021 actggaacccgacacatcgaaatttgcgaacacacacatacttggtaagtggctaccccaa 1080
 Db 1021 actggaacccgacacatcgaaatttgcgaacacacacatacttggtaagtggctaccccaa 1080
 Qy 1081 aacgatctgcttggctaccgcgatgaccggtgcctttatcacccatgctggttcccatggt 1140
 Db 1081 aacgatctgcttggctaccgcgatgaccggtgcctttatcacccatgctggttcccatggt 1140
 Qy 1141 gtttatgaagcatatgcaatggcttcccatggtgatgatgcccttgggttgatcacg 1200
 Db 1141 gtttatgaagcatatgcaatggcttcccatggtgatgatgcccttgggttgatcacg 1200
 Qy 1201 atggacaatgcaaaagcgcgtgagacgaaggaagcgtgaggtgacccgaatgttctgaaa 1260
 Db 1201 atggacaatgcaaaagcgcgtgagacgaaggaagcgtgaggtgacccgaatgttctgaaa 1260
 Qy 1261 atgactctgaagatttagaaatgctctaaagcagtgatcaatgacaaaagtacaaag 1320
 Db 1261 atgactctgaagatttagaaatgctctaaagcagtgatcaatgacaaaagtacaaag 1320
 Qy 1321 gagaacatcatgagcctctccagccttcacaagaccgcgcggtgagccgctggacctg 1380
 Db 1321 gagaacatcatgagcctctccagccttcacaagaccgcgcggtgagccgctggacctg 1380
 Qy 1381 gccgtgtcttgggtgagttgtgatgagcacaagggcgccacacccctgcgcccgcga 1440
 Db 1381 gccgtgtcttgggtgagttgtgatgagcacaagggcgccacacccctgcgcccgcga 1440
 Qy 1441 gccacgacctccacctggtaccagttaccatctcttgacgtgattggtttcctctggcc 1500
 Db 1441 gccacgacctccacctggtaccagttaccatctcttgacgtgattggtttcctctggcc 1500
 Qy 1501 gtcgtgctgacagtgcccttcacaccttttaaatgttgtgttatgctacccgaaatgc 1560
 Db 1501 gtcgtgctgacagtgcccttcacaccttttaaatgttgtgttatgctacccgaaatgc 1560
 Qy 1561 ttggggaaaaagggcagtttaagaaagcccaataatccaagaccattgagaagtgggt 1620
 Db 1561 ttggggaaaaagggcagtttaagaaagcccaataatccaagaccattgagaagtgggt 1620
 Qy 1621 gggaataaaggttaaaatttgaacattccctcctagtcatttccaaacttgaaaacagaatc 1680
 Db 1621 gggaataaaggttaaaatttgaacattccctcctagtcatttccaaacttgaaaacagaatc 1680
 Qy 1681 agtgttaaatctattttatttatttaaggaataactttgataataataatcagccccag 1740
 Db 1681 agtgttaaatctattttatttatttaaggaataactttgataataataatcagccccag 1740
 Qy 1741 agtgccttaaaaaattctttaaataaaaaataatagactcgtcagtaagaatattt 1800
 Db 1741 agtgccttaaaaaattctttaaataaaaaataatagactcgtcagtaagaatattt 1800
 Qy 1801 gaatatgtatcgtgccccctccggtgtctttgatcaggtgacatgtgccatttttcaga 1860
 Db 1801 gaatatgtatcgtgccccctccggtgtctttgatcaggtgacatgtgccatttttcaga 1860

Qy 1861 ggacgtgcagacaggtgcgcttctagattacttttctactctgaaacatggccctgttt 1920
 Db 1861 ggacgtgcagacaggtgcgcttctagattacttttctactctgaaacatggccctgttt 1920
 Qy 1921 ggagatgcggattcaaaagtggttcccaacgctgcacctactgcaaaatggcagttttaa 1980
 Db 1921 ggagatgcggattcaaaagtggttcccaacgctgcacctactgcaaaatggcagttttaa 1980
 Qy 1981 ctatcttttggcttctgcagatggttgcaattgaccttaaccataatgggtcagtcct 2040
 Db 1981 ctatcttttggcttctgcagatggttgcaattgaccttaaccataatgggtcagtcct 2040
 Qy 2041 catctctctgctctcatagtgccaccttgggttttaaaagaagggaagcctttgtacc 2100
 Db 2041 catctctctgctctcatagtgccaccttgggttttaaaagaagggaagcctttgtacc 2100
 Qy 2101 tttagagtgtagtgaaatgaatggctggctgagtcgactgagaacacatgatgttt 2160
 Db 2101 tttagagtgtagtgaaatgaatggctggctgagtcgactgagaacacatgatgttt 2160
 Qy 2161 ctgcttttgggaaaaaagaatgactatgaaattgggtgggtgtatttgagaagata 2220
 Db 2161 ctgcttttgggaaaaaagaatgactatgaaattgggtgggtgtatttgagaagata 2220
 Qy 2221 atcattgcttatgtcaaatggagctgaatttgataaaaccccaaaacacagctatgaagt 2280
 Db 2221 atcattgcttatgtcaaatggagctgaatttgataaaaccccaaaacacagctatgaagt 2280
 Qy 2281 gctgggcaaatgttactttttcttctgctgatttctcacaact 2320
 Db 2281 gctgggcaaatgttactttttcttctgctgatttctcacaact 2320

RESULT 3

AA027369
 ID AAQ27369 standard; cDNA; 2351 BP.
 XX

AC AAQ27369;
 XX

DT 27-JAN-1993 (first entry)
 XX

DE HUG-Br1.
 XX

XX Billirubin; UDP-glucuronosyltransferase; HUGBr1; HUGBr2;
 monoglucuronide; diglucuronide; ss.
 XX

OS Homo sapiens.
 XX

PH Key Location/Qualifiers
 FT CDS 16..784
 FT /tag= a
 FT polyA_signal 2330..2335
 FT /tag= b
 FT /number= 1
 FT polyA_signal 2338..2343
 FT /tag= c
 FT /number= 2
 XX

PN W09212987-A.
 XX

PD 06-AUG-1992.
 XX

PF 10-JAN-1992; 92WO-US00282.
 XX

PR 10-JAN-1991; 91US-0639453.
 XX

XX (USSH) US DEPT HEALTH & HUMAN SERVICE.
 PA Owens IS, Ritter JK;
 XX

XX WPI; 1992-284593/34.
 DR P-PSDB; RAR26153.
 XX

XX WPI; 2002-075093/10.
 DR P-PSDB: ABB04429.
 XX
 PT Combinations of flavopiridol and an agent that increases conjugative
 PT enzyme activity or glucuronosyltransferase activity, with reduced side
 PT effects, for treating cancer -
 XX
 PS Claim 80: Page 137-140; 145pp; English.
 XX
 CC The present invention relates to a method of reducing the toxicity of
 CC flavopiridol by administration in combination with a second agent that
 CC increases conjugative enzyme activity or glucuronosyltransferase
 CC activity. This second agent should be capable of inhibiting biliary
 CC transport and may be a uridine 5'-diphosphate glucuronyltransferase such
 CC as that encoded by the present sequence (designated UGRIA9). The method
 CC can be used in the treatment of cancer, gastrointestinal diseases and
 CC parasitic diseases.
 XX
 SQ Sequence :320 BP; 608 A; 495 C; 547 G; 670 T; 0 other;

Query Match 100.0%; Score 2320; DB 24; Length 2320;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 2320; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ctcaagctgcagttctctgatggctgcacaggggtgacccagccctctctctatgtgtg 60
 DB 1 ctcaagctgcagttctctgatggctgcacaggggtgacccagccctctctatgtgtg 60

QY 61 tct 120
 DB 61 tct 120

QY 121 gggagccactggttcacatcagctgctggtggtggaacactcattctcagggggcatgag 180
 DB 121 gggagccactggttcacatcagctgctggtggtggaacactcattctcagggggcatgag 180

QY 181 ctggttct 240
 DB 181 ctggttct 240

QY 241 aagacttattcaactctatcaccctggaggtgctggccgggggttcaggctttgtcc 300
 DB 241 aagacttattcaactctatcaccctggaggtgctggccgggggttcaggctttgtcc 300

QY 301 catgtctcaatggaagcacaagttacgaagtatatattctctatttaagtgtcacaact 360
 DB 301 catgtctcaatggaagcacaagttacgaagtatatattctctatttaagtgtcacaact 360

QY 361 gacattttgacttatttttttcaaatgaggtttgtttaaagacaaaaaattagta 420
 DB 361 gacattttgacttatttttttcaaatgaggtttgtttaaagacaaaaaattagta 420

QY 421 gaactactaaagagaggtcttttgaagcaggttctctcgatctcttctgataactgtg 480
 DB 421 gaactactaaagagaggtcttttgaagcaggttctctcgatctcttctgataactgtg 480

QY 481 ttaattgttgcataattct 540
 DB 481 ttaattgttgcataattct 540

QY 541 cactatcttgaagaaggtgcaggtgcccctgctctctctctctctctctctctctctctct 600
 DB 541 cactatcttgaagaaggtgcaggtgcccctgctctctctctctctctctctctctctctct 600

QY 601 tttaggttctcagatgcacttctcaagagagagtagtacggaaccacatcatgtcacttg 660
 DB 601 tttaggttctcagatgcacttctcaagagagagtagtacggaaccacatcatgtcacttg 660

QY 661 gaggacatttattatgccacgttttttcaaaaatgcctagaaaatagcctctgaaatt 720
 DB 661 gaggacatttattatgccacgttttttcaaaaatgcctagaaaatagcctctgaaatt 720

QY 721 ctcaaaacactgttacgagtagtctctacagccacacatcaaatgttggtgtgca 780
 DB 721 ctcaaaacactgttacgagtagtctctacagccacacatcaaatgttggtgtgca 780

QY 781 acgagcttctgttttgagcatctccaaacccggtgagcccaacatgattcttctgtggt 840
 DB 781 acgagcttctgttttgagcatctccaaacccggtgagcccaacatgattcttctgtggt 840

QY 841 atcaactgcacacaggaagccgttgcctatggaattggaagcctacataatctctct 900
 DB 841 atcaactgcacacaggaagccgttgcctatggaattggaagcctacataatctctct 900

QY 901 ggaacacatggaattgtgtttctctctctctctctctctctctctctctctctctctct 960
 DB 901 ggaacacatggaattgtgtttctctctctctctctctctctctctctctctctctctct 960

QY 961 aaagctatggaattgctgattgtttgggcaaaatccctcagacagctctctgtggcggtac 1020
 DB 961 aaagctatggaattgctgattgtttgggcaaaatccctcagacagctctctgtggcggtac 1020

QY 1021 actggaacccgacacatcgaattcttgggaacacacagatacttgaagtgcctaccccaa 1080
 DB 1021 actggaacccgacacatcgaattcttgggaacacacagatacttgaagtgcctaccccaa 1080

QY 1081 aacgatctgttggtccacccgatgacccgtgctcttatacccatgctggttcccatggt 1140
 DB 1081 aacgatctgttggtccacccgatgacccgtgctcttatacccatgctggttcccatggt 1140

QY 1141 gtttatgaagcatatgcaatggcgttcccatggtgagtgatgagccctgttgggtgatcag 1200
 DB 1141 gtttatgaagcatatgcaatggcgttcccatggtgagtgatgagccctgttgggtgatcag 1200

QY 1201 atggacaatgcacacgcatggaagctggaagctggaagtgacccctgaatgttcttgaaa 1260
 DB 1201 atggacaatgcacacgcatggaagctggaagctggaagtgacccctgaatgttcttgaaa 1260

QY 1261 atgacttctgaagatttagaaaaatgctctaaagcagtcataatgacaaaaagtacaa 1320
 DB 1261 atgacttctgaagatttagaaaaatgctctaaagcagtcataatgacaaaaagtacaa 1320

QY 1321 gagaacatctgcgctctccagcttccacagggaccgccgggtggagccgctggaactg 1380
 DB 1321 gagaacatctgcgctctccagcttccacagggaccgccgggtggagccgctggaactg 1380

QY 1381 gccgtgtctgggtgagttgtgtaggacacagggcgccacacactgcccgcgca 1440
 DB 1381 gccgtgtctgggtgagttgtgtaggacacagggcgccacacactgcccgcgca 1440

QY 1441 gccac 1500
 DB 1441 gccac 1500

QY 1501 gtcgtgtcagagtgagcttctacacacacacacacacacacacacacacacacacacac 1560
 DB 1501 gtcgtgtcagagtgagcttctacacacacacacacacacacacacacacacacacacac 1560

QY 1561 ttgggggaaaaagggcgaggttaagaaagcccaacacacacacacacacacacacacacac 1620
 DB 1561 ttgggggaaaaagggcgaggttaagaaagcccaacacacacacacacacacacacacacac 1620

QY 1621 ggggaaataaggttaaaatttgaacacacacacacacacacacacacacacacacacacac 1680
 DB 1621 ggggaaataaggttaaaatttgaacacacacacacacacacacacacacacacacacacac 1680

QY 1681 agtgttaaaatct 1740
 DB 1681 agtgttaaaatct 1740

QY 1741 agtgtttaaataatct 1800
 DB 1741 agtgtttaaataatct 1800

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 21, 2002, 06:34:42 ; Search time 682.45 Seconds
(without alignments)
5836.679 Million cell updates/sec

Title: US-09-835-082-1
Perfect score: 2320
Sequence: 1 ctacagtcagttctctgat.....ttctgatgttctctacaact 2320

Scoring table: IDENTIFY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 1736436 seqs, #58457221 residues

Total number of hits satisfying chosen parameters: 3472872

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N: Geneseq_032802.*

1:	/SIDSL/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.*
2:	/SIDSL/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.*
3:	/SIDSL/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.*
4:	/SIDSL/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.*
5:	/SIDSL/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.*
6:	/SIDSL/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.*
7:	/SIDSL/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.*
8:	/SIDSL/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.*
9:	/SIDSL/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.*
10:	/SIDSL/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.*
11:	/SIDSL/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.*
12:	/SIDSL/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.*
13:	/SIDSL/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.*
14:	/SIDSL/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.*
15:	/SIDSL/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.*
16:	/SIDSL/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.*
17:	/SIDSL/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.*
18:	/SIDSL/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.*
19:	/SIDSL/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.*
20:	/SIDSL/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.*
21:	/SIDSL/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.*
22:	/SIDSL/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.*
23:	/SIDSL/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.*
24:	/SIDSL/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	2320	100.0	2320	24	ABA05198 Human uridine 5'di
2	2320	100.0	2320	24	ABA05199 Human uridine 5'di
3	1581.4	68.2	2351	13	AAQ27369 HUG-Brl. Homo sap
4	1576.8	68.0	2368	13	AAQ27370 HUG-Brl. Homo sap
5	1551.8	66.9	1655	23	AAS91252 DNA encoding novel
6	1004.4	43.3	18887	24	AAS18543 DNA encoding UDP g
7	998.6	43.0	1001	21	AAH51403 Human UGT1A7 relat
8	870.6	37.5	1001	21	AAH51404 Human UGT1A7 relat
9	798	34.4	1667	13	AAQ33027 Exon 5 from the US

10	769.6	33.2	1001	21	AAH51405 Human UGT1A7 relat
11	766.6	33.0	930	21	AAZ45116 UDP-glucuronosyltr
12	741.2	31.9	951	21	AAZ45115 UDP-glucuronosyltr
13	735	31.7	735	21	AAZ45118 UDP-glucuronosyltr
14	659.8	28.4	759	21	AAZ45117 UDP-glucuronosyltr
15	578.8	24.9	874	24	ABL01468 Murine apoptosis f
16	364.2	15.7	391	20	AAZ07208 Human lung tumour
17	364.2	15.7	391	21	AAC79066 Human lung tumour
18	364.2	15.7	391	23	AAD23141 Human UGT1A7 relat
19	359	15.5	1001	21	AAH51406 Human UGT1A7 relat
20	353	15.2	1001	21	AAH51407 Human UDP-glucuron
21	345.2	14.9	1854	21	AAZ95200 DNA encoding novel
22	345	14.9	1766	23	AAZ91380 Uridine diphospho
23	344.4	14.8	2107	19	AAV15900 Human colon cancer
24	342	14.7	380	21	AAC98723 Human UDP-glucuron
25	339.8	14.6	2092	21	AAZ95199 Human UDP-glucuron
26	333.2	14.4	1976	21	AAZ95206 Human carbohydrate
27	332.8	14.3	1650	21	AAC65396 DNA encoding novel
28	327.2	14.1	2802	23	AAS69711 Human drug metabol
29	325.6	14.0	2111	23	AAS69712 Human breast cance
30	322.4	13.9	1636	24	AAD24666 Human DNA encoding
31	322.4	13.9	1722	22	AAS46185 Human breast cance
32	322.4	13.9	2974	22	AAS69710 Human colon cancer
33	318	13.7	1859	23	AAD06844 Human UGT1A7 relat
34	314.6	13.6	1713	22	AAS91253 DNA encoding novel
35	302.4	13.0	508	21	AAC98478 Rat hepatocyte car
36	296	12.8	1001	21	AAH51408 Exons 3-4 from the
37	282	12.2	2025	23	AAS91253 DNA encoding novel
38	230.4	9.9	366	21	AAH87085 Rat hepatocyte car
39	222.4	9.6	775	13	AAQ33026 DNA encoding novel
40	220.8	9.5	273	21	AAH87491 DNA encoding novel
C 41	203	8.8	2448	23	AAS74393 DNA encoding novel
C 42	203	8.8	2448	23	AAS74393 DNA encoding novel
C 43	203	8.8	2448	23	AAS74393 DNA encoding novel
C 44	203	8.8	2448	23	AAS74393 DNA encoding novel
C 45	203	8.8	2448	23	AAS74393 DNA encoding novel

ALIGNMENTS

RESULT 1
ABA05198
ID ABA05198 standard; cDNA; 2320 BP.
XX ABA05198;
AC ABA05198;
XX
DT 04-MAR-2002 (first entry)
DE Human uridine 5'diphosphate glucuronyltransferase UGT1A9 cDNA #1.
XX
XX Human; uridine 5'diphosphate glucuronyltransferase enzyme; UGT1A9;
KW flavopiridol; cancer; gastrointestinal disease; parasitic infection;
KW cytostatic; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 19..1611
FT FT /*tag= a
FT FT /product= "UGT1A9"
XX
XX WO200180896-A2.
XX
XX 01-NOV-2001.
XX
XX 12-APR-2001; 2001WO-US12526.
XX
XX HUG-Brl. Homo sap
PR 21-APR-2000; 2000US-0553829.
XX
XX (ARCH-) ARCH DEV CORP.
XX
XX Retain MJ, Innocenti F, Iyer L;

Search completed: August 21, 2002, 11:59:18
Job time: 32806 sec

Oy		gtagctaccaccaaacagatcgttggtcacccgatgcacccgttcetttatccccatgc	1127
Ddb		GTGGTCCACAGCTTTACCATCGGAAACTTCAGCACACGCCGGTGTCATCACTCATTTG	1128

TYPE: NUCLEIC ACID
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cdna
PCT-US92-00282-18

APPLICANT: RITTER, JOSEPH K.
TITLE OF INVENTION: THE GENETIC LOCUS UGT1 AND A MUTATION THEREIN.
NUMBER OF SEQUENCES: 40
CORRESPONDENCE ADDRESS:
ADDRESSEE: CUSHMAN DABRY & CUSHMAN
STREET: 1615 L STREET, N.W.
CITY: WASHINGTON
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20036-5601

Query Match 7.0%; Score 161.8; DB 5; Length 1190;
Best Local Similarity 51.8%; Pred. No. 1.3e-36;
Matches 414; Conservative 0; Mismatches 382; Indels 3; Gaps 2;

```

QY 68 tgcctgacgtgtgcttccagagcaggaagactactgttagtgccttgatggagacc 127
D 140 TGCTGGCCCACTGTGTCCCACTGCTGGGAAGATGCTGTGATCCCACTGGATGGCACC 199
QY 128 actggttcaaccatgaggtcggtggtgagaaactcattctcagggggcaggtggtg 187
D 200 ACTGGCTGAGCATGCTTGGGGCCATCCAGCAGCTGCAGAGAGGGGACATGAATAGTTG 259
QY 188 tagtcagccagagtggtggtggtggtggtggtggtggtggtggtggtggtggtg 247
D 260 TCCTAGCACCTGACGCTGCTGTGTACATCAGAGACGGAGCATTTTACACCTTGAAGCGT 319
QY 248 attcaacttcaataccctgaggtggtggtggtggtggtggtggtggtggtggtggtg 305
D 320 ACCCTGTGCCAATCCAAAGGAGGATGTGAAGAGTCTTTTGTGTAGTCTCGGGCATATG 379
QY 306 tcaatggaagcacaagtaacgaagtatatattctcttataatggttgcatacaatgacat 365
D 380 TTTTGTAGATGATCTTCTCGCAGCGTGTATCAAAACATACAGAAATAAAGG 439
QY 366 tttgact-tattttttcaaatgaggtgtgttttaagaacaaataattagtagaat 424
D 440 ACTGTGTGCTTTTGTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGT 499
QY 425 acttaagagaggttcttttgcagaggttctctgcagcttctctgcagcttctctgcag 484
D 500 CCTGGCAGAAAGACGCTTTCATGTCATGCTGACGACCCCTTCTCTCTCTGACGCCCA 559
QY 485 ttgttgcataattctctctctctctctctctctctctctctctctctctctctctct 544
D 560 TCGTGGCCCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 619
QY 545 atcttgaagaggtgcaagtgctctctctctctctctctctctctctctctctctctct 604
D 620 TGGAAATTTGAGGTACCCAGTGCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 679
QY 605 gttctcagatgcatgactttcaagagagagagagagagagagagagagagagagagag 664
D 680 CTCATTCAGATCAGATGACCTTCTGCGAGGGGTGAAGACATGCTCATTTGCTTTTAC 739
QY 665 aacattattatgccaccgttttttcaaaaatgccctagaaaagcctctgaaattctcc 724
D 740 AGAATTTCTGTGCGACGTGTTTATTTCCCGTATGCAACCTTGCTCTCAGAAATTCCTTC 799
QY 725 aacacctgttaccgagtagatctctacagccacacataaatttgggttggagacagg 784
D 800 AGAGAGAGGTGACTGTGCAGACCTATGAGCTCTGCTGCTGCTGCTGCTGCTGCTGCTG 859
QY 785 actttgtttggactatcccaaacccgtgtagtgcacacatgatctctcattggtgtagta 844
D 860 ACTTTGTGAAGATGACCTAGGCCATATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 919
QY 845 actgcatcaggaagacc 863
D 920 ACTGCTTCCACCAATACC 938

```

RESULT 10

PCT-US92-00282-12
Sequence 12, Application PC/TUS9200282
GENERAL INFORMATION:
APPLICANT: OWENS, IDA S.

```

Query Match 6.9%; Score 161.2; DB 5; Length 1197;
Best Local Similarity 52.6%; Pred. No. 2e-36;
Matches 421; Conservative 0; Mismatches 373; Indels 6; Gaps 3;

QY 85 gcagagcaggaagactactgttagtgccttgatggagaccactgg--ttcaccatga 142
D 304 GCTGAGAGTGAAGAGTGTGTTGGTGGTCCCATGATGCGACGCCCTGTGCTCAGTCAATG 363
QY 143 ggtcgtgtgtggaagaaactattctcagggggcaggtggtggtggtggtggtggtggt 202
D 364 GGGAGCCCTTGGGGAGGCTCCATGCCAGAGGCCACAGCGGTGGTCTCACCACGAGG 423
QY 203 tgagttggcaactgggaagatcactgaattgcacagtgaagacttattcaacttcatata 262
D 424 TGAATATGCATCAAAAGAGAGAAATTTTTCACCTGACAGGCTATGCTGTTCATGGA 483
QY 263 ccttgagagatcgtgacggaggttcaagcttttggccatgctcaat---ggaagcac 319
D 484 CCCAGAGGAAGTATGATGCTGTACCTGGCTACACCTGAGGGGTCTTTGAAACAGAAC 543
QY 320 aagtaagaagtatatattctcttataatgggttcaacaatgacatttttgcacttttt 379
D 544 ATCTTCTGAAGAGATATTTAGAAAGTATGCAATTTATGAACAATGATCTTTGGCCCTTC 603
QY 380 tttaaatgicagaggtttgttttaagacacaaataattagtaataacttaagagaggtt 439
D 604 ATAGGTGTTGTGGAGCTACTGCATAATGAGGCCCTGATCAGGCACCTGAATGCTACTT 663
QY 440 cttttagtcagtggtttctcgtatcccttttgaataactggtggtttaaattgttgccaaat 499
D 664 CCITTGATGTGTTTAAACAGACCCCGTTAACCTCTGTGGGGGGGTGCTGGCTAAGTACC 723
QY 500 tctcctccctccctggtctctcgtcaggggaatactttgcccactatcttgaagaaggtg 559
D 724 TGTCGATTCCTGCTGTGTTTTTTTGGAGGTATCCATGCTAGACTTTAGAGCCA 783

```



```
Db 303 GGTGAATATGTACATCAAGAGAGAACTTTTTCACCCCTGACAAAGTATGCCATTTTCATG 362
Qy 261 tacctggagagatcgaccggaggttcaaggcttttgccatgtctcaat---ggaaagc 317
Db 363 GACCCAGGACGAATTTGATCGCCTTTTGTGGGTCAACACATCAATGTTTGTGAACAGA 422
Qy 318 acaagtaagaatatactctatttaattggtttacatacaatgacatttttgactatt 377
Db 423 ACATCTTCGTAGGAATTTCTAGAGAAATGGAATATGCAATATGCTTTTGATCAT 482
Qy 378 ttttcaaatgagaggtttgtttaagacaaaataatagtagataacttaagagagag 437
Db 483 ACATAGGTCTTGTGTGGAGTACTGCAATAGAGGCCCTGATCAGGCACCTGCATGTAC 542
Qy 438 tcttttgatgaggtttctcgatcttttgataactgtggttaattgttgccaaata 497
Db 543 TTCTTTGATGTGTTCTACAGACCCCTTTCACCTCTGCGCGGGTGTGCTAGTA 602
Qy 498 tttctccctccctccgtggtcttccgagggggaataactttgcccactatcttgaaag 557
Db 603 CCTGTGATTCCTGTGTGTTTCTTGAGGAACATTTCCATGTGATTTAGACTTTAAAGG 662
Qy 558 tgacagtgccctgctctcttcttatgttcccccagaaattctcttaggggttccagatgc 617
Db 663 CACACAGTCTCCAAACCTTCCCTCTATATCTTAGATTACTAAGACCAATTCAGACCA 722
Qy 618 catgactttcaagagagatgacggaacacatcatgcacttgagagaaactttattatg 677
Db 723 CATGACATTCCTGCAAGGGTCAAGAACATGCTCTACCCCTGTCGCTACCTTTG 782
Qy 678 ccaccgttttttcaaaaatgcccctagaaatagcctctgaaattctccaaacacctgttac 737
Db 763 CCATGCTGTTCTGCTCTTATGCAAGCCCTTGCCCTCTGAGCTTTTTCAGAGAGAGGTC 842
Qy 738 ggagatgatctctacagcacacacatcaaatgttgggttggaagagactttgttggga 797
Db 843 AGTGTGGATCTTGTGACCCATGATCCTGTGTGGGTGTTCGAGGGAGTTGTGATGGA 902
Qy 798 ctatcccaaacccgtgagtcaccaatgatcttcttattggttggttatcaactgcatcagg 857
Db 903 TTACCCAGGCGCATGTCGCCAACATGCTCTTCATTGGGGGCATCAACTGTGCGCAACGG 962
Qy 858 aaagccgtgacctgaagaattgaagcctacatta 892
Db 963 GAAGCCACTATCTCAGGTCGTATTTGGTGCCTTTA 997
```

RESULT 7

```
; Sequence 16, Application PC/TUS9200282
; GENERAL INFORMATION:
; APPLICANT: OWENS, IDA S.
; TITLE OF INVENTION: THE GENETIC LOCUS UGT1 AND A MUTATION
; TITLE OF INVENTION: THEREIN.
; NUMBER OF SEQUENCES: 40
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CUSHMAN DAREY & CUSHMAN
; STREET: 1615 L STREET, N.W.
; CITY: WASHINGTON
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20036-5601
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US92/00282
; FILING DATE: 19920110
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
```

```
; NAME: SCOTT, WATSON T.
; REGISTRATION NUMBER: 26581
; REFERENCE/DOCKET NUMBER: 91532-PC
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-861-3000
; TELEFAX: 202-822-0944
; TELEX: 6714627 CUSH
; INFORMATION FOR SEQ ID NO: 16:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1448 base pairs
; TYPE: NUCLEIC ACID
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: YES
; PCT-US92-00282-16
```

```
Query Match 8.0%; Score 186.6; DB 5; Length 1448;
Best Local Similarity 53.8%; Pred. No. 1.1e-43;
Matches 429; Conservative 0; Mismatches 364; Indels 2;

Qy 81 ctttgcaggagcaggagactgtagtgcctatgtagtggagagccactgtttccaccat 140
Db 178 CTGGGCTGAGAGTGGGAAGGTGCTGTGTGTACCCATTCATGTCGAGCCACTGGCTCG-CAT 236
Qy 141 gaagtcggtggtgagaaactcattctcaggggcagatgagtggttagtcatgcagaga 200
Db 237 GCGGAGGCGGTGCGGAGGTCTCCATGCCAGAGCCACAGCGGTGGTCTTCACCCGAGA 296
Qy 201 ggtgagttggcaactgggaagatcaactgacagtggaagacttattcaacttcata 260
Db 297 GGTGAATATGCATCAACAGAGAAATTTTTCACCCCTGACAACTATGCACTTTTCATG 356
Qy 261 tacctggagagatcgaccggaggttcaaggctttgcccagctcctcaat---ggaaagc 317
Db 357 GACCCAGGATGAATTTGATGCTTGTGCTGGCCACACTCAACTGTACTTGTGAACAGA 416
Qy 318 acaagtaagaatatactctctatttaattggtttggtttacatacaatgacatttttgactatt 377
Db 417 ACATTTTCTGAGACATTTTCTAAAAGTATGCAATTTTGAATAATTCATCTTTGGTCTT 476
Qy 378 ttttcaaatgacagaggtttgttttaagacaaaaaattagtagaatactataagagagag 437
Db 477 TCATAGTCTTGTGTGAGGTACTGCATTAATGAGGCCCTGATCAGGCACCTGATGTAC 536
Qy 438 tcttttgatgcagtggtttctcgatccttttgataactgtggttggttaattgttgcacaa 497
Db 537 TTCCTTCGATGTGGTTTAAATAGACCCAAATTTACCTCTGTGGGCGAGTCTGGCTAAGTA 596
Qy 498 tttctccctccctccgtggttcttcagcaggggaatactttgccactatcttgagagag 557
Db 597 CCTATCAATTCCTGCTGTGTTTTTTTGTGAGNACATTCATGTGATTTAGACTTTAAGG 656
Qy 558 tgcagagtcctcgtcctcttcttctatgtccccaagaattctcttaggtttctcagatgc 617
Db 657 CACACAGTGTCCAAATCCTTCTCCTCTATATTCTTAAGTTACTAACGACCAATTCAGACCA 716
Qy 618 catgactttcaagagagagtagcagccacatcatgcacttgcaggaacatttattatg 677
Db 717 CATGACATTCCTGCNAAGGGTCAAGAACATGCTCTACCCCTCTGCCCCCTGCTCCTACCTTTG 776
Qy 678 ccaccgttttttcaaaaatgcccctagaaatagcctctgaaattctccaaacacctgttac 737
Db 777 CCATGCTCTTCTGCTCTTATGCAAGCCCTTGCCCTCTGAGCTTTTTCAGAGGAGGTGTC 836
Qy 738 ggagatgatctctacagccacacatcaatttgggttggtgcaacggactttgttttggga 797
Db 837 AGTGTGGATCTTCTCAGCCATGCATCTGTGTGGGTGTTCGAGGGGACTTGTGTGATGGA 896
Qy 798 ctatcccaaacccgtgtagtgcaccaatgatctctaatgttggttatcaactgcatcagg 857
Db 897 TTACCCAGGCGCATGTCGCCAACATGCTTTCATTGGGGGCATCAACTGTGCGCAACAG 956
```



```

QY 1505 tctgacagtgcccttcaccccttaaaatgtgtgtcttattggtacccggaatgctgg 1564
Db 1496 TGCTGACAGTGGCTTCATCACCTTTAAATGTTGTCTTATGGCTACCGGAAATGCTTGG 1555
QY 1565 ggaataaagggcagtgtaagaaagccacaaatccaaagaccattgagaagtggtggga 1624
Db 1556 GGAAAAAGGGCGAGTTAGAAAGCCACAAATCCAAAGACCCTTGAGAGTGGGTGGGA 1615
QY 1625 aataaggtataaatttgaaacattcccttagctatttccaaacttgaaacagaaatcagtg 1684
Db 1616 AATAAGGTAAATTTTAAACCATTCCTTAGTCTATTTCCAACTTCCAAACAGAAATCAGTG 1675
QY 1685 ttaaatctatttattcttattaaagaaatcatttgcataaaatcattcagcccccagagtg 1744
Db 1676 TTAATATCATTTTATTCTTATTAGGAATACCTTTGCATAAATTAATACGCCCCAGAGTG 1735
QY 1745 ctttaaaaaattcttataaaaaataatagactcgttagctagtcagtaagaattttgaat 1804
Db 1736 CTTTAAAAAATCTCTTAAATAAAATAATAGACTCGCTAGTCAATTAAGATATTGTAAT 1795
QY 1805 atgtatcgtgccccctccggtgtcttcttgatcaggtgacatgtgccatttttcagaggac 1864
Db 1796 ATGTATCTGCCCCCTCCGGTCTCTTTGATCAGGATGACATGSCCATTTTTCAGAGGAC 1855
QY 1865 gtgcagacaggtggcattctagattcttcttactctgaacatggcctgtttggga 1924
Db 1856 GTGCAGACAGGCTGGCATTTCTAGATTACTTTTCTTACTCTGAAACATGGCCTGTTGGGA 1915
QY 1925 gtgcgggattcaaaagtggtccaccgctgcccctactgcaaatggcagttttaaactta 1984
Db 1916 GTGCGGATTCAAAGGTGTCCACCGCTGCCCTACTGCAATGGCAGTTTAACTTTA 1975
QY 1985 tctttggctctgcagatggttgcatttgcatttcccttaaccataatggctcctcatc 2044
Db 1976 TCTTTGGCTCTGCGAGATGSGTGGCAATGATCCTTTAAACCAATTAATGGTCAGTCCATC 2035
QY 2045 tctgtcctcttccataggtggccacctgtgtgttttaaaaggaagcctttgaccttta 2104
Db 2036 TCTGTCTCTGCTTCATAGTGGCCACCTGTGTGTTTAAAGAGGGAAGCTTTGTACCTTTA 2095
QY 2105 gagttaggtgaaatgaatgaatgctgagtgagtgacactgagacagcatatgattcttg 2164
Db 2096 GAGTGTAGTGAATGAATGAATGGCTTGGAGTGCACATGAGAACAGCATATGATTTCTTG 2155
QY 2165 cttbgggaaaaagaatgatgctatgaaattggtggtggtgtatttgagaagataatca 2224
Db 2156 CTTTGGGAAAAAGAAATGATGCTATGAAATTTGGTGGTGGTGTATTTTGAGAAGATAATCA 2215
QY 2225 ttgcttatgcaaatggagctgaatttgataaaaaaccacaaatacacagctatgaagtgctg 2284
Db 2216 TTGCTTATGTCNAATGGAGCTGAATTTGATTAATAAACCCCAAAATACAGCTATGAGTGCTG 2275
QY 2285 ggaagtttacttttttcttgatgatttcttccacaact 2320
Db 2276 GGCAAGTTTACTTTTTTCTGATGTTTCTCTACAAC 2311

```

RESULT 3

PCT-US92-00282-25

; Sequence 25, Application PC/US9200282

; GENERAL INFORMATION:

; APPLICANT: OWENS, IDA S.

; APPLICANT: RITTER, JOSEPH K.

; TITLE OF INVENTION: THE GENETIC LOCUS UGT1 AND A MUTATION

; TITLE OF INVENTION: THEREIN.

; NUMBER OF SEQUENCES: 40

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: CUSHMAN DARBY & CUSHMAN

; STREET: 1615 L STREET, N.W.

; CITY: WASHINGTON

; STATE: D.C.

; COUNTRY: U.S.A.

```

; ZIP: 20036-5601
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US92/00282
; FILING DATE: 19920110
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: SCOTT, WATSON T.
; REGISTRATION NUMBER: 26581
; REFERENCE/DOCKET NUMBER: 91532-PCT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-861-3000
; TELEFAX: 202-822-0944
; TELEX: 6714627 CUSH
; INFORMATION FOR SEQ ID NO: 25:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1561 base pairs
; TYPE: NUCLEIC ACID
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; PCT-US92-00282-25

```

Query Match 41.7%; Score 967.2; DB 5; Length 1561;

Best Local Similarity 98.0%; Pred. No. 3.3e-268;

Matches 1011; Conservative 0; Mismatches 18; Indels 3; Gaps 3;

```

QY 1291 aaagcagtcctcaatgacaaaaagttaacaaaggagaacataatgcctctccagcctcaac 1350
Db 248 AACCAAACTGCTCTTTGTGTAGTTTACAAAGAGAACATCATGCGCTCTCCAGCCTTAC 307
QY 1351 aaggaccgcccggtagcgccgtgacccgtgtctgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1410
Db 308 AAGGACGCCCGGTGGAGCGCGTGACCTGGCCGTGTTCTGGGTGGAGTTTGTGTAGG 367
QY 1411 cacaaagcgccgcacacacctgcccgcagccacacctcacctggtaccagttaccat 1470
Db 368 CACAAGGGCGCCACACCTGGCGCCGACGCCACGACCTCACCTGTTACCAAGTACCAT 427
QY 1471 tccctgacgtgattgggttctcttctgtgcccgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1529
Db 428 TCCTTGGACGTGATTGGTTTCTCTTGGCGGTCTGGCTGGACAGTGGCTTACACCTT 487
QY 1530 taaatgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1589
Db 488 TAAATGTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 547
QY 1590 ccacaaatccaaagacccttgagaaagtgggtgggaaataaggttaaaatttttgaaacattc 1649
Db 548 CCACAAATCCAAAGACCCATTGAGAAGTGGGTGGGAAATAAGTAAATTTTGAACCATTC 607
QY 1650 cctagtcatctccaaacttgaaacagaaatcagtgtaaatcatttttatttatttattt 1709
Db 608 CCTAGTCATTTCCAAACTTGAACACAGAAATCAGTGTAAATTTTATTCTATTATTAAAG 667
QY 1710 gaaatacttgcataaataatcagccccagagtgctttaaaaaattctcttaataaaaa 1769
Db 668 GAAATACTTTGCATAAATAATCAAGCCGCCAGAGTGGCTTTAAAAAATTTCTCTTAAATAAAA 727
QY 1770 aataagactcgttagtaagataatttgaatgtatgtgtgtgtgtgtgtgtgtgtgtgtgt 1828
Db 728 ATAATAGACTCGGTAGTACGTAAAGATATTTGAATATGTATGTGTGTGTGTGTGTGTGTGT 787
QY 1829 ttgatcagatgacatgtgcccatttttccagagagcgtgcagacagcgtggcattctaga 1888
Db 788 TTTGATCAGGATGACATGTGCCATTTTTCAGAGGACGTGCAGACAGAGGTGG-ATTCTAGA 846
QY 1889 ttacttttctactctgaaaaaatgcccgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1948

```


Qy 141 gaggcggtggtggaagaaactcattctcagggggcatgagggtggttgatcatgccaag 200
Dy 132 GCGGAGGCGCTTCGGGAGCTCCATGCGCAGAGGCCACCGAGGGGTGGTCTCCACCCCGA 191
Qy 201 ggtgagttggaactgggaagactcaactgaaatgacagtgaaagacttatcaactica 260
Dy 192 GGTGAATATGACATCAAGAGAGAGAAATTTTCACCTGACAGCCTATGCTGTCCATG 251
Qy 261 taccctggagatctggacccggaggtcaggggttcaagggttttggccatgctcaat---ggaagc 317
Dy 252 GACCAGAGGAATTTGATCGGGTTACGCTTACGCTGGCTACACTCAAGGGTTCTTTGAAACAGA 311
Qy 318 acaagtcagaaatataatctctcattatggttgcatacaatgacatltttgacttatt 377
Dy 312 ACATCTTCTGAGAGATATCTAGAGTATGGCAATTTGACAAATGATCTTTGGCCCT 371
Qy 378 ttttccaaatgcaggagttgttttaagagaaaaaataatagtagaataacttaagagagag 437
Dy 372 TCATAGGTGTTGTGGAGCTACTGCATAATAGAGGCCCTGATCAGACCCCTGAATGCTAC 431
Qy 438 tttctttgatgcaggtgtttctcgatccttttgataactgtgcttaattgttgcataa 497
Dy 432 TTCTTTGATGTGGTTTAAACAGACCCGTTAACTCTCGGGGGGTGCTGGCTAAGTA 491
Qy 498 tttctccctccctcccggtggttcgcaagggaataactttgccaactatcttgaagagg 557
Dy 492 CCGTTCGATTCCTGCTGCTGTTTTTTTGGAGGTACATTCGATGTGACTTAGACTTTAAGG 551
Qy 558 tgacagtgccgtgctcct 617
Dy 552 CACA-CAGTGTCCAAATCCTCTCTCTATATCTTAACTTACTTAAGACCAATTCAGACCA 611
Qy 618 catgacttcaaggagagtagcaggaaccacatcatgactctggaggaacatttatatg 677
Dy 612 CATGACATCTCGCAAGGGTCAAGAACATGCTCTACCTCTCGGCCCTGCTTACATATG 671
Qy 678 ccacgggttttcaaaaaagccttagaagtagcctctgaaatctcccaaacacactgttac 737
Dy 672 CCATACTTTTTCTGCCCCCTATGCAAGTCTTGCTCTGAGCTTTTCAGAGAGAGGTGTC 731
Qy 738 ggaatgatctctacagcacacatacaattgttgggttggaagcagacttttttggga 797
Dy 732 AGTGTGATCTTGTGAGTATGATCGTGTGGCTGTGCTTCCAGAGGGACTTTGTGATGA 791
Qy 798 ctatcccaaacccgtgatcccaacatgatctctcattgtgtgatacgaactgcaatcagg 857
Dy 792 CTAC-CCAGGCGCATCATGCCAACATGGTCTTCAATTTGGGGCATCAACTGTGCCAACGG 851
Qy 858 aaagcsggtgcctatgaaatttgaagcctacattaatgctctctggagaaacatggaattg 917
Dy 852 GAAG-CACTATCTCAGGAATTTGAAGCCTACATTAATGCTTCTGAGAACATGGAATGT 911
Qy 918 ggtttctctttggatcaatggttcagaaattccagagaagaagcctatggaactgttc 977
Dy 912 GGTTTTCTCTTTGGGATCAATGGTCTCAGAAATTCAGAGAAAGATATGCAATTC 971
Qy 978 tgatctttgggcaaaatccctcagacagtcctgtggcggtacactggaacccagaccatc 1037
Dy 972 TGAT-CTTTGGGCAAAATCCCTCAGACAGTCCTGTGGCGGTACATGGAACCCGACCATC 1031
Qy 1038 gaattctgcgaacacacgataactgtttaagtggctaccaccaaacgactctgttgta 1097
Dy 1032 GAAT-CTTGGCAACAAACGATACTTTGTAAGTGGCTTACCCCAAAACGATCTGCTTGTCA 1091
Qy 1098 cccgatgaccgtgctttatcaccatgctgttcccatggtgtttatgaagacatg 1157
Dy 1092 CCGATGACCGTGCCTTTATCACCCTATGCTGGTTCCTCCATGCTGTTATGAAGCATATG 1151
Qy 1158 caatggcgttcccatggtgatgacctgtttgtgtgtatcagatggacaatgcaagcg 1217
Dy 1152 CAATGCGCTTCCATGGTGATGCCCTGTTGTTGGTGTATGATGAGTGAACAATGCAAGCG 1211

Qy 1218 catgagagataaggagctgagtgacccctgaatgttctggaataactctctgaagattt 1277
Dy 1212 CATGGAGACTAAGGGAGCTGGAGTGCACCTGAATGTTCTTGGAAATGACTTCTGAAGATTT 1271
Qy 1278 agaaaaatgctctaaaaagcagtcataatgacaaagttaacaaggagaaacatcatcgccct 1337
Dy 1272 AGAAATGCTCTTAAAGCAGTCAATCAATGACAAAGTTTACAAAGAGAACATCAGGCGCT 1331
Qy 1338 ctcaagccttcaaaaggaacccggggggggcggctgagacctgagccgtgtcttgggtgga 1397
Dy 1332 CTCCAGCCTTCAACAAGAGCCGCCGGTGGAGCGCTGGACCTGGCCGTCTTGGGTGGA 1391
Qy 1398 gtttgtatgagcaacaaggcgccgacacactcgcccccagccagccacacacacactcg 1457
Dy 1392 GTTGTGATGAGGCAACAAGGGCGGCCACACCTCGCCCGCCGACCCACACCTCACCTG 1451
Qy 1458 gtaccagtaacattctctggacgtgattggtttctctcttggccgtgctgctgacagtgc 1517
Dy 1452 GTACCAGTACCATTCTTGGACGTGATTTGCTTGGCCGTCTGCTGGTGCAGATGCG 1511
Qy 1518 ctcatcacctttaaattgtgtgttatggtctacccgaaatgcttgggggaaaaaaggcg 1577
Dy 1512 CTTCATCACCTTTAAATGTTGTGCTTATGGCTACCGGAAATGCTTGGGAAAAAGGGG 1571
Qy 1578 agttaagaagcccaacaaatccaagaccattgagaagtgggtgggaaaaaagaagtaaat 1637
Dy 1572 AGTTAAGAAAGGCCACAAATCCAAGACCCATTTGAGAAAGTGGTGGGAAATAGGTAAAT 1631
Qy 1638 ttgaaacattcctctagtcatttccaaacttgaacagaaatcagtggttaataatctttt 1697
Dy 1632 TTTGAACCATCTCCTAGTCAATTTCCAACTTGAACAGAACTAGTAAATTCATTTT 1691
Qy 1698 attcttataagaataactcttgcaataataatcaagcccaagagctcttaaaaaattc 1757
Dy 1692 ATCTTATTAAGGAAATACTTTGCAATAAATAATCAGCCCCAGAGTCTCTTAAAAAATTC 1751
Qy 1758 tcttaaaataaaataatagactcgtcagtaaaagataattgaaatgatatctctgccc 1817
Dy 1752 TCTTAATAAATAAATAATAGCTAGTACGCTAGTCAAGTAAAGATATTTGAATATGATCGTCCC 1811
Qy 1818 cctcgggtgcttttgatagatgacatgcccatttttccagagagcgtcagacagct 1877
Dy 1812 CTTCCGGTGTCTTTGATCAGGATGACATGTGCCATTTTTCAGAGAGCTGCACACAGCT 1871
Qy 1878 ggcattctagattactttcttactctgaaacatggcctgttttgggagtcggtggaatcaa 1937
Dy 1872 GGCATTCTAGATTACTTTTCTTACTCTGAAACATGGCCTGTTGGAGTGGCGGATCAA 1931
Qy 1938 agtgtgtccacgcgtccctcactgcaaatggcagtttataatcttcttctgtctct 1997
Dy 1932 AGGTGCTCCACGCTGCCCTTACTGCAAAATGGCAGTTTTTAATCTTATCTTTTGGCTCT 1991
Qy 1998 gcagatggttgcaattgactcttaaccaataatggtcagtcctcatctctgtcctcttc 2057
Dy 1992 GCAGATGTTGCAATGATCTCTTAACCAATAATGTCAGTCTCTCTCTCTCTCTCTCT 2051
Qy 2058 atagttaccactgtgtttttaaagaaggagctttgacacttttagagtgtagtcaa 2117
Dy 2052 ATAGTGGCCACCTTGTGTTTAAAGAGGGAAGCTTTGTACCTTTTAGAGTGTAGTGAA 2111
Qy 2118 atgaatgaattgggtgagtcactggaacacacatgatttcttcttgggggaaaaa 2177
Dy 2112 ATGAATGAATGGCTTGGAGTGCACCTGAGAACAGACATATGATTTCTTCTTGGGAAAAA 2171
Qy 2178 gaatgacttatgaataatggtgggtggttatattgagaagataaactcttctctatcaa 2237
Dy 2172 GAATGATGCTATGAATATGGTGGGTGGTGTATTTGAGAAGATATCATTCATTCATA 2231
Qy 2238 atggagctgaatttgataaaaaaaccccaaaatacacagctatgaagtcgtgggcaagttaactt 2297
Dy 2232 ATGGAGCTGAATTTGATAAAAAACCCAAAAATACAGCTATGAAGTGTGGCAAGTTACTT 2291
Qy 2298 tttttctgagtttctcacaact 2320

ORIGIN

Query Match 27.8%; Score 644; DB 10; Length 723;
 Best Local Similarity 96.7%; Pred. No. 1.1e-157;
 Matches 700; Conservative 0; Mismatches 20; Indels 4; Gaps 4;

QY 1078 caaagcagatctgttgatccacgagatgacccggtccctttatccaccatgctggttcccat 1137
 Db 2 CAAGAGCATCTCTGTGTCACCGGATGACCGTG-CTTTATCATCCCATGCTGGTTCGCAT 60
 QY 1138 ggtttattgaagcatatgaatggcgttcccatggtgagtgatgagtccttgttggatg 1197
 Db 61 GGTGTTATGAAGCATATGCATGGCGTTCCTCATGATGATGCCCTGTGTTGGTAT 120
 QY 1198 cagatggacaatgcgaagcgcgcatggagactaaggagctggagctgagctgacctaagtcttg 1257
 Db 121 CAGATGGACAATGCAAGCGCATGGAGACTAAGGGAGCTGGAGTGACCTGAATGTTCTG 180
 QY 1258 gaattgactctgaagattgaaaaatgctctaaaagcagtcacatcaatgacaaaagtac 1317
 Db 181 GAATGACTCTCTGAAGATTTAGAAATGCTCTAAAAGCAGTCATCAATGACAAAAGTTAC 240
 QY 1318 aagagaacatcatcgctctccacgcttcacagccttcacagacgcgcgcgtggagccgctggac 1377
 Db 241 AAGAGAACATCATCGCTCTCCAGCCTTTCACAAAGGACGCCCGGTGGAGCGCTGGAC 300
 QY 1378 ctggcgcgtgtctgtgggtgagttgtgatgagcacaagggcgccacacacctgcgccc 1437
 Db 301 CTGGCGGTGTCTGTGGTGGAGTTGTGATGAGGCACAAAGGCGGCCACACCTGGGCCCC 360
 QY 1438 gcagccacagacactcaactggaccagctacacttccctggagctgattggttccctcttg 1497
 Db 361 GCACCCACGACCTCACCTGGTACCACTACCATCTCTGGACGTGATGGTTCTCTCTG 420
 QY 1498 gccgtgctgacagtggtccttcacacttcaactttaaattgtgtctatgctacccgaaa 1557
 Db 421 GCCGTCTGTGTGACAGTGGCCTTCATCACCTTTAAATGTTGTGCTTATGGCTACCGGAAA 480
 QY 1558 tgcctgggaaaaaggcaggttaagaaagcccaaaatccaaagccattgagaagt 1617
 Db 481 TGCTTGGGAAAAGGCGAGTTAGAAAGCCCAAAATCCAAAGACCCCTTGAGAAGTG 540
 QY 1618 ggtgggaaataaggttaaaa-ttttgaaaccattccctagctatccattccaaacttgaaaaacag 1676
 Db 541 GGTGGGAAATAAGTAAATTTTGAACCAATTCCTAGTCATTTCCAAACTTTGAAACAG 600
 QY 1677 aatcagtgtaaatcttttattcttatttaaggaataactttgcatataaataatcagcc 1736
 Db 601 AATCAGTGTACATTCATTTTATTTTATTAAGGAATAC-TTGCAATAATTAATCAG-C 658
 QY 1737 ccagagtgtttaaaaaattctotataaaaaataatagactcgctagtcagtaagaat 1796
 Db 659 CCAGAGTGCTTTAAAAAATTTCTTAATAAATAATATAGACTCGTAGTCAGTAAGATTGAT 718
 QY 1797 attt 1800
 Db 719 ATGT 722

RESULT 14

LOCUS BG469506 829 bp mRNA linear EST 21-MAR-2001
 DEFINITION 602532818F1 NIH_MGC_15 Homo sapiens CDNA clone IMAGE:4660540 5',
 mRNA sequence.
 ACCESSION BG469506
 VERSION BG469506.1 GI:13401781
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 829)

AUTHORS
TITLE
JOURNAL
COMMENT

NIH-MGC http://mgc.nci.nih.gov/.
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-r@mail.nih.gov
 Tissue Procurement: ATCC
 CDNA Library Preparation: Ling Hong/Rubin Laboratory
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: NIH Intramural Sequencing Center
 Clone Distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: LLC1458 row: c column: 05
 High quality sequence stop: 795.

FEATURES

source

1..829
 Location/Qualifiers
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:4660540"
 /clone_lib="NIH_MGC_15"
 /tissue_type="adenocarcinoma cell line"
 /lab_host="DH10B (phage-resistant)"
 /note="Organ: colon; Vector: pOTB7; Site: 1: XhoI; Site: 2:
 EcoRI; cDNA made by oligo-dT priming. Directionally
 cloned into EcoRI/XhoI sites using the following 5'
 adaptor: GGCACGAG(G). Size-selected >500bp for average
 insert size 1.8kb. Library constructed by Ling Hong in
 the laboratory of Gerald M. Rubin (University of
 California, Berkeley) using ZAP-cDNA synthesis kit
 (Stratagene) and Superscript II RT (Life Technologies)"

BASE COUNT 186 a 190 c 207 g 246 t

ORIGIN

Query Match 27.8%; Score 641.2; DB 10; Length 829;
 Best Local Similarity 89.9%; Pred. No. 6.4e-157;
 Matches 732; Conservative 0; Mismatches 78; Indels 4; Gaps 4;

QY 5 gctgcagttctctgatggttcacaggggtggaccagccctctctctctatgtgtgtc 64
 Db 18 GCTGCAGTTCCTCATGGCTGCGCAGGGTGGACAGCCCGTTCCTTTATGTGTGTGTC 77
 QY 65 tgcgtgcacctgtggttgcgcagagcaggaagactactgtagtgcacatgagtgga 124
 Db 78 TACTGTGACCTGTGGCTTGGCCGAGGAGGAAAGCTGCTGGTAGTGCCTATGATGGGA 137
 QY 125 gccactggttcaccatgaggtcggtggtggaagaaactcattctcagggggcagtggtg 184
 Db 138 GTCACGTGTTACCATGTCAGTGGTGGTGGAGAAACTTATCTCTCAGGGGCATGAGTGG 197
 QY 185 ttgtagtcacagaggtgagttggcaactgggaagatcactgaattgcacagtgaaga 244
 Db 198 TTGTAGTCAATGCCAGAGGTGAGTTGGCAACTGGAAAGATCACGAATTCACAGTGAAGA 257
 QY 245 ctattcaactctcatataccctggaggtcctgaccggaggttcaaggcttttggccatg 304
 Db 258 CTTACTCAACCTCGTACACTCTGGAGATCAGAACCGGGAATTCATGTTTTCGCCCATG 317
 QY 305 ctcaatggaaagcagagtcacagaggttgcctctcttatttaagagcaaaaattagtagaat 364
 Db 318 CTCAATGGAAAGCAGAGGCACAAAGATATATTTTCTCTATTAATGAGTTTCATCCAGTGGTT 377
 QY 365 ttttgactatttttttcaaatgcagaggttgcctttaaagagcaaaaattagtagaat 424
 Db 378 TTCTTGACTTATTTTTCGATTCAGGAGTTGGTTTATATGACCCGAAAATTTAGTAGAAT 437
 QY 425 acttaagagagaggttcttttgatgcaggtgttctc-tcgatccttttgaataactgtggtta 483
 Db 438 ACTTAAAGGAGAGAGTCTCTTTGGATGCAGTGTTCGTGGATCTCTTTGATACCTGTGGCTTA 497
 QY 484 attgtgcacaaatatttctcctccctccctccgtggtcttcgcagcagggggaataactttgccac 543
 Db 498 ATTGTGCTGAATATTTCTCCCTCCCTCTCTGTGGTCTTCACCGAGGGGAATATTTCGCCAC 557


```

QY 683 gtttttcaaaatgccctag--aaatagcctctgaattctcc-aaacacctgttacgg 739
Db 660 GTTTTTTCAAAAATGCCCTAGGAAATAGCCCTCTGAAATNCTCCAAAACACCTGTACGG 719
QY 740 agtatgctctacagcacacatcaatttggtgtgtgcgaacggactttgtttgaact 799
Db 720 AGTATGATCTCTACAGCACACATCAACTCGGCTCGTGGCAGCGACTTGTCTTGACT 779
QY 800 atccaaaacccgtgatgcccaacatgatcttcatttggtggtatcaac--tgccatcaggg 857
Db 780 AT-CCAAACCCGTGATGCCACAAATGATCCTCTCTGGGGGTATCCACCGCGCTTACGG 838
QY 858 aaacgcgttgctatgaattga 881
Db 839 AAACCCGTGCTATGGATTGGA 862

RESULT 8
AI307269/c
LOCUS AI307269 818 bp mRNA linear EST 08-APR-1999
DEFINITION tb24d12 x1 NCI_CGAP_kid12 Homo sapiens cDNA clone IMAGE:2055287 3'
similar to gb:M57899 UDP-GLUCURONOSYLTRANSFERASE 1A PRECURSOR,
MICROSOMAL (HUMAN);, mRNA sequence.
ACCESSION AI307269
VERSION AI307269.1 GI:4002458
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
NCI-CCGAP http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgaps-r@mail.nih.gov
Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
Emmert-Buck, M.D., Ph.D.
cDNA Library Preparation: M. Bento Soares, Ph.D.
cDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CCGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html
Seq primer: -40UP from Gibco
High quality sequence stop: 453.
FEATURES
source
location/Qualifiers
1...818
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:2055287"
/clone.lib="NCI_CGAP_Kid12"
/tissue_type="2 pooled tumors (clear cell type)"
/lab_host="DH10B"
/notes="Organ: kidney; Vector: pMT3D-Pac (Pharmacia) with
a modified polylinker; Site:1: Not 1; Site:2: Eco RI;
Plasmid DNA from the normalized library NCI_CGAP_Kid5 was
prepared, and ss circles were made in vitro. Following HAP
purification, this DNA was used as tracer in a subtractive
hybridization reaction. The driver was PCR-amplified cDNAs
from a pool of 5,000 clones made from the same library
(cloneIDs 1323912-1325831, 1471368-1472903 and
1492104-1493255). Subtraction by Bento Soares and M.
Fatima Bonaldo."
BASE COUNT 247 a 177 c 134 g 258 t 2 others
ORIGIN

Query Match 30.0%; Score 695.8; DB 9; Length 818;
Best Local Similarity 95.9%; Pred. No. 3e-171;
Matches 745; Conservative 0; Mismatches 29; Indels 3; Gaps 3;

```

```

QY 1547 gctaccggaatgtttggggaaaaaggcg-agttaagaagcccaaatccaaagacc 1605
Db 818 GCTACGAAAATGCTGGGGGAAAAAGGCGAATTTAAGNAGGCCNCAATCCCAAGAAC 759
QY 1606 cattagaagtggtgggaataaagtaaaatttgaaacattccctcattcatcaaa 1665
Db 758 ATTTGAGAAGTGGGTGGGAATAAAGGTAAATTTGAACCAATCCCAAGTCAATTCCTCCAA 699
QY 1666 ctggaacacagaatca-gtgttaaatcattttatttatttaagaagaataactttgata 1724
Db 698 CTGGAACAGAAATCAGGGTAAATTCATTTATTTCTTTAAGGAATATCTTGGGATA 639
QY 1725 aattaatcagcccc-agatgctttaaaaaattctcttaataaaaaataatagactcgct 1783
Db 638 AATTAATCAGCCCCACAGAGTCTTTTAAANATTCCTTTAATAAATAAATAATAGACTCGCT 579
QY 1784 agtcagtaagatatattgaatatgtatogtgcocccctccgggtgtttttgatcagatgac 1843
Db 578 AGTCAGTAAAGATATTTGAATATGTATCGTGCCTCCCTCTGGTCTTTTGATCAGAGTAC 519
QY 1844 atgtccattttcagaggcgtgcagacagcgtggtcattctagattacttttttactc 1903
Db 518 ATGTGCATTTTTCAGAGGACGTGCAGACAGCGCTGGCAATCTAGATTACTTTTCTTACTC 459
QY 1904 tgaacatgacctgtttgggagtgccgaggttcaaaagtggtcccccacctgcccactctg 1963
Db 458 TGAACATGCGCTGTTGGGAGTGGCGGATTCAAAGGTGGTCCACGCTGCCCTACTG 399
QY 1964 caaatgcagtttttaactttatcttttggcttcgcagatggttgcaattgatcttaac 2023
Db 398 CAATGGCAGTTTAACTTATCTTTTGGCTTCGAGATGGTGGCAATGATGATCTTAAAC 339
QY 2024 caataatgctcagtcctcctctctctctctctctctctctctctctctctctctctct 2083
Db 338 CAATAATGCTCAGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 279
QY 2084 aagggaagcttgctacatttagatgtagtgaatgaatgaatgaatgagcttgagtgactg 2143
Db 278 AAGGGAAGCTTTGACCTTTAGAGTGTAGTGTAGTGTAGTGTAGTGTAGTGTAGTGTAGT 219
QY 2144 agaacgcgatgatgtttctgtttgtttgtttgtttgtttgtttgtttgtttgtttgtttgt 2203
Db 218 AGAACGACATATGATTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCT 159
QY 2204 gtgtattgagaagaataatcattgttctcgaatgagctgaattgtataaaaccca 2263
Db 158 GTGTATTTCAGAAGATAATCATTTGTTCAATGGAGCTGAATTTGATATAAAACCCA 99
QY 2264 aaatacagctatgaagtgctggcgaatttttctctctctctctctctctctctctct 2320
Db 98 AAATACAGCTATGAAGTCTGGGCAAGTTTACTTTTTTCTGATGTTTCTCTACAACT 42

```

```

RESULT 9
BG289777
LOCUS BG289777 881 bp mRNA linear EST 21-FEB-2001
DEFINITION 602385071F1 NIH_MGC_93 Homo sapiens cDNA clone IMAGE:4514129 5',
mRNA sequence.
ACCESSION BG289777
VERSION BG289777.1 GI:13045907
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 881)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgaps-r@mail.nih.gov
Tissue Procurement: ATCC

```


found through the I.M.A.G.E. Consortium/LLNL at:

http://image.llnl.gov

Plate: LLAM10407 row: m column: 12

High quality sequence stop: 746.

FEATURES

source

1..882

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="IMAGE:4516403"

/clone_lib="NIH_MGC_93"

/tissue_type="transitional cell papilloma, cell line"

/lab_host="DH10B (phage-resistant)"

/note="Organ: bladder; Vector: pCMV-SPORT6; Site:1: NotI;

Site:2: SalI; Cloned unidirectionally; oligo-dT primed.

Average insert size 1.7 kb. Library enriched for

full-length clones and constructed by Life Technologies.

Note: this is a NIH_MGC Library."

259 a 155 c 202 g 266 t

BASE COUNT

ORIGIN

Query Match 31.18; Score 722.4; DB 10; Length 882;

Best Local Similarity 94.5%; Pred. No. 3.3e-178;

Matches 792; Conservative 0; Mismatches 41; Indels 5; Gaps 4;

QY 1447 gacccacgtgtaccagtcaccattccttgagcgtgattggttctctcttccctgcccgtg 1506

DB 1 GACCTCACCCTGGTACCAGTACCATTCTTGGACGTGATTGGTTCTCTTGGCCGTCGTG 60

QY 1507 ctgcaatgacctcattcatttaattgtcttattgcttaccggaatgcttggg 1566

DB 61 CTGACATGGCCCTCATCACCCTTAAATGTTGTGCTTATGGCTACCGGAATGCTGGG 120

QY 1567 aaaaaggcgagtgaaagagcccaatccaaagaccatgagaagtgagtgaggaa 1626

DB 121 AAAAAGGGGAGTTAAGAAAGGCCAACAAATCCAAAGCCATTGAGAAGTGGTGGAAA 180

QY 1627 taagtaaaatttgaccattcccttagctattccaaacttgaaacagaaatcagttt 1686

DB 181 TAAGTAAATTTTGAACCATTCCTTAGTCAATTTCCAACTTGAACAGAAATCAGTGT 240

QY 1687 aaatcattttattcttatttaagaaatactttgtcataaataatcagcccgagtgct 1746

DB 241 AAATTCATTTTATTCTTATTAAAGAAATACCTTGGCATAAATTAATCAGCCAGAGTCT 300

QY 1747 taaaaaattctttaaataaaataagactcgtactgactgactgactgactgactgact 1806

DB 301 TTAATAAATTCCTTAAATAAATAATAGACTCGCTAGCTAGCTAGCTAGCTAGCTAG 360

QY 1807 gtatggtcccccctcggtgtctttgtatcagatgacatgtgccattttccagagcgt 1866

DB 361 GTATCGTCCCGCTCGGCTGTCTTTGATCAGGATGACATGTGCCATTTTCAGAGGAGT 420

QY 1867 gacagcggctgacatctagattactttcttactctgaaacatgacctgtttggagtt 1926

DB 421 GCAGACAGGTGGCATCTAGATTACTTTTCTTACTTGAACATGGCCTGTTGGAGT 480

QY 1927 gggggattcaagggtgtccaccgctgcccctactgcaaatggcaggttttaattttatc 1986

DB 481 GCGGGATTCAAGAGTGTGCCACCGCTGCCCTTACTGCAAAATGGCAGTTTAACTTATC 540

QY 1987 ttttgctctcagatggttgcaatgatccttaaccaataatggctcagtcctcatc 2046

DB 541 TTTTGGCTTCTGCAGATGGTTGCAATGTGATCCCTTAACCAATAATGGTCACTCCTCATCTC 600

QY 2047 tgcctgcttca tagtgccacctgtgt-gtttaagaagggaagctttgtacctttag 2105

DB 601 TGCTCGCTTCATAGGTGCGACCTTGTGTGGTTTAAAGAAGGGAAGCTTTGTACCTTTAG 660

QY 2106 agtgtaggtg-aatgaatgaatggttgagtgacgtgagacagcatatgatctt-t 2163

DB 661 AGTGTAGGTGAAATGAATGATGTTGGAGGTGCACGTGAGAACAGCATATGATTTCTGG 720

QY 2164 gctttgggaaagaatgatgctgaattgtaaaaccccaatacacagctatgaagtg 2281

DB 721 GCTTTGGGAAAAAGAAATGATGCTATGGAAATGGGGGTGGGATTCGAGAAAAATATC 780

QY 2224 atgcttatgtaaatgagctgaattgataaaaccccaatacacagctatgaagtg 2281

DB 781 ATGCTTATGTCRAAGGAGG--TGATTTGATTAACCAAAATACAGGTATTAAGGTG 836

RESULT 6

BI759297

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

human.

Human.

Human.

Human.

Human.

Human.

Human.

Human.

Human.

Human.

Human.

Human.

Human.

Human.

Human.

Human.

Human.

Human.

Human.

Human.

Human.

Human.

Human.

Human.

Human.

Human.

Human.

Human.

Human.

Human.

Human.

Human.

Human.

Human.

Human.

Human.

Human.

Human.

Human.

Human.

Human.

Human.

Human.

Human.

Human.

Human.

Human.

Human.

Human.

Human.

Human.

Human.

Human.

Human.

Human.

Human.

Human.

Human.

Human.

Human.

Human.

Human.

QY 2164 gctttgggaaagaatgatgctgaattgtaaaaccccaatacacagctatgaagtg 2281

DB 721 GCTTTGGGAAAAAGAAATGATGCTATGGAAATGGGGGTGGGATTCGAGAAAAATATC 780

QY 2224 atgcttatgtaaatgagctgaattgataaaaccccaatacacagctatgaagtg 2281

DB 781 ATGCTTATGTCRAAGGAGG--TGATTTGATTAACCAAAATACAGGTATTAAGGTG 836

BI759297

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

human.

Human.

Human.

Human.

Human.

Human.

Human.

Human.

Human.

Human.

Human.

Human.

Human.

Human.

Human.

Human.

Human.

Human.

Human.

Human.

Human.

Human.

Human.

Human.

Human.

Human.

Human.

Human.

Human.

Human.

Human.

Human.

Human.

Human.

Human.

Human.

Human.

Human.

Human.

Human.

Human.

Human.

Human.

Human.

Human.

Human.

Human.

Human.

Human.

Human.

Human.

Human.

Human.

Human.

Human.

Human.

Human.

Human.

Human.

Human.

Human.

Human.

QY 2164 gctttgggaaagaatgatgctgaattgtaaaaccccaatacacagctatgaagtg 2281

DB 721 GCTTTGGGAAAAAGAAATGATGCTATGGAAATGGGGGTGGGATTCGAGAAAAATATC 780

QY 2224 atgcttatgtaaatgagctgaattgataaaaccccaatacacagctatgaagtg 2281

DB 781 ATGCTTATGTCRAAGGAGG--TGATTTGATTAACCAAAATACAGGTATTAAGGTG 836

BI759297

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

human.

Human.

Human.

Human.

Human.

Human.

Human.

Human.

Human.

Human.

Human.

Human.

Human.

Human.

Human.

Human.

Human.

Human.

Human.

Human.

Human.

Human.

Human.

Human.

Human.

Human.

Human.

Human.

Human.

Human.

Human.

Human.

Human.

Human.

Human.

Human.

Human.

Human.

Human.

Human.

Human.

Human.

Human.

Human.

Human.

Human.

Human.

Human.

Human.

Human.

Human.

Human.

Human.

Human.

Human.

Human.

Human.

Human.

Human.

Human.

Human.

Human.

QY 2164 gctttgggaaagaatgatgctgaattgtaaaaccccaatacacagctatgaagtg 2281

DB 721 GCTTTGGGAAAAAGAAATGATGCTATGGAAATGGGGGTGGGATTCGAGAAAAATATC 780

QY 2224 atgcttatgtaaatgagctgaattgataaaaccccaatacacagctatgaagtg 2281

DB 781 ATGCTTATGTCRAAGGAGG--TGATTTGATTAACCAAAATACAGGTATTAAGGTG 836

BI759297

LOCUS


```

QY 1073 taacccaaacagatcgtctgtggtcaccgcagatgacccgtgccttattatccaccatcgtggtt 1132
Db 1180 TACCCCAAAATGATCTGCTTGGTCATCCAAAGACTCGGCGCATTCATCAACACACTCTGGCT 1239
QY 1133 cccatggtgtttatgaagacatgcaatggcgttcccatggtgatgagccctgtgttg 1192
Db 1240 CCCATGGTATTATGAAGGAATATGCAATGGAGTTCCGATGGTGATGATGCCCTATTG 1299
QY 1193 gtctcagatgacaatgcaaacgcagatggagactaaaggagcgtgagtgaccctgaatg 1252
Db 1300 GCGTCAGATGAGCAATGCAACGCGCATGGAACTCGGGAGGTGGGTGACCCCTGAATG 1359
QY 1253 ttctggaatgactcttgaagatttagaaaatgctctaaaagcagtcacatcaatgacaaaa 1312
Db 1360 TCCTTGAATGACTGCTGATGATTGGAAATGCCCCTTAAACTGTCTATCAACAACAAGA 1419
QY 1313 gtacaaagaaacatcatcgccctctcagccttcacaggaagccgcgggtgagccgc 1372
Db 1420 GCTACAAGGAACATCATCGCCCTCTCAGCCCTTCAAGAGACCGTCTATAGAGCCTC 1479
QY 1373 tggacctggcgtgtctggtggagtttgtgatgagcagcagcgcgcacacctgc 1432
Db 1480 TGGACCTGCTGCTGCTGGGTGGAATACGTGATGAGGCACAAGGGGGCACACACCTGC 1539
QY 1433 gccccgcagccacagacctcaacctggtaccagtgaccatccttgagcgtgattggttcc 1492
Db 1540 GCCCGCGCCCGCCATGACCTCACCTGGTATCATGATACCTCCTTGGATGATTGGCTTCC 1599
QY 1493 tcttgccgtgctgctgacagtggtgcttcacaccttaaatgttgttcttatggtacc 1552
Db 1600 TCCTGGCCATGCTGTTCACAGTGGTTCATTGCTTTAAATGTTGCTGCTATGGCTGCC 1659
QY 1553 ggaatgtctggggaacaaaggcagtgtaagaagcccaacaaatcccaagaccattgag 1612
Db 1660 GGAATGCTTTGGGGGAAGGGCGGAGTGAAGAAATCACACAAATCCCAAGACCCATTGAG 1719
QY 1613 aagtgggtgggaataaggtaaatgttgaaaccttcccttagctcattttccaaactgaaa 1672
Db 1720 AAGTGGGGGAAGTGAAGGAGAAGTATTAGTTTCAATTCATCTGATGATGTTGAACCTGGA 1777
QY 1673 acagatcagtgtaaatcatttattcttattaaagaaatcatttcataaattatc 1732
Db 1778 ----ACAAAGTGTAGTATCTGATGCTGTTGTTAGGGAATAATTCACCATACATTATA 1833
QY 1733 agccccagatgcttataaaatctcttta--aataaaataatagacgtcgtagtcagt 1790
Db 1834 CATTCAGCACATTTAAATAATAATAATAAATCTAAATGCTGGCCACACCCATCAGG 1893
QY 1791 aaagattttgaatgatctgctgccccctccgtgctcttcttgatcagatgacatgccc 1850
Db 1894 GAAGTCTAGTATATCTGATGCTGCTGCTTTCCAGTATCTTCAGTCTAGACAACTCTTGCC 1952
QY 1851 attttcagagagcgtgcagacagcgtggcattcttagatt 1890
Db 1953 ATCTGT-TGGTAATTTACAGAAAGTCTGGCAGTCTGCTTT 1991

RESULT 3
AK002629 2202 bp mRNA linear HTC 19-JAN-2002
LOCUS Mus musculus adult male kidney cDNA, RIKEN full-length enriched
DEFINITION library, clone:0610013P07:UDP-glucuronosyltransferase 1 family,
member 1, full insert sequence.
ACCESSION AK002629
VERSION AK002629.1 GI:12832753
KEYWORDS HTC; CAP trapper.
SOURCE Mus musculus (strain: C57BL/6J) adult male kidney cDNA to mRNA,
clone:lib:RIKEN full-length enriched mouse cDNA library
Clone:0610013P07.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (sites)

```

AUTHORS TITLE JOURNAL MEDLINE PUBMED REFERENCE AUTHORS

TITLE JOURNAL MEDLINE PUBMED REFERENCE AUTHORS

TITLE JOURNAL MEDLINE PUBMED REFERENCE AUTHORS

TITLE JOURNAL MEDLINE PUBMED REFERENCE AUTHORS

TITLE JOURNAL

COMMENT

FEATURES source

Carninci,P. and Hayashizaki,Y.
High-efficiency full-length cDNA cloning
Meth. Enzymol. 303, 19-44 (1999)
99279253
10349636
2 (sites)
Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K.,
Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new genes
Genome Res. 10 (10), 1617-1630 (2000)
20499374
11042159
3 (sites)
Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P.,
Konno,H., Akiyama,J., Nishi,K., Kitsuai,T., Tashiro,H., Itoh,M.,
Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishine,T., Harada,A.,
Yamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K.,
Fujiwaki,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watahiki,M.,
Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuura,S., Kawai,J.,
Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.
RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multicapillary sequencer
Genome Res. 10 (11), 1757-1771 (2000)
20530913
11076861
4 (sites)
The RIKEN Genome Exploration Research Group Phase II Team and the
FANTOM Consortium.
Functional annotation of a full-length mouse cDNA collection
Nature 409, 685-690 (2001)
5 (bases 1 to 2202)

Adachi,J., Aizawa,K., Akahira,S., Akimura,T., Aono,H., Arai,A.,
Arakawa,T., Baldarelli,R., Bono,H., Brownstein,M., Bult,C.,
Carninci,P., Fukuda,S., Fukunishi,Y., Furuno,M., Hanaoka,T.,
Hara,A., Hayatsu,N., Hill,D., Hiramoto,K., Hiraoka,T., Hori,F.,
Hume,D., Imatani,K., Ishii,Y., Itoh,M., Izawa,M., Kasukawa,T.,
Kato,H., Kawai,J., Kojima,Y., Konno,H., Kouda,M., Koya,S.,
Kurihara,C., Matsuyama,T., Miyazaki,A., Nishi,K., Nomura,K.,
Numazaki,R., Ohno,M., Okazaki,Y., Okido,T., Owa,C., Quackenbush,J.,
Saito,H., Saito,R., Sakai,C., Sakai,K., Sano,H., Sasaki,D.,
Schriml,L., Shibata,K., Shibata,Y., Shinagawa,A., Shiraki,T.,
Sogabe,Y., Suzuki,H., Tagami,M., Tagawa,A., Takahashi,F.,
Tanaka,T., Tejima,Y., Toya,T., Yamamura,T., Yamanaka,I.,
Yasunishi,A., Yoshida,K., Yoshino,M., Muramatsu,M. and
Hayashizaki,Y.

Direct Submission
Submitted (10-JUL-2000) Yoshihide Hayashizaki, The Institute of
Physical and Chemical Research (RIKEN), Laboratory for Genome
Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
Kanagawa 230-0045, Japan (E-mail:genome-res@gsc.riken.go.jp,
URL:http://genome.gsc.riken.go.jp/, Tel:81-45-503-9222,
Fax:81-45-503-9216)
Please visit our web site (<http://genome.gsc.riken.go.jp/>) for
further details.

cDNA library was prepared and sequenced in Mouse Genome
Encyclopedia Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in RIKEN.
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues. First strand cDNA was primed with a primer
[5' GAGAGAGAGCGCCGCACTCGAGTGTGTGTGTGTGTGTGTGT 3'], cDNA was
prepared by using trehalose thermo-activated reverse transcriptase
and subsequently enriched for full-length by cap-trapper. cDNA went
through one round of normalization to Rot = 5.0. Second strand cDNA
was prepared with the primer adapter of sequence[5'
GAGAGAGAGCGCCGCACTCGAGTGTGTGTGTGTGTGTGTGT 3']. cDNA was
cleaved with XhoI and SstI. Cloning sites, 5' end: SstI; 3' end:
XhoI. Host: SOLR.

Location/Qualifiers
1. .2202
/organism="Mus musculus"
/strain="C57BL/6J"

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 21, 2002, 02:43:12 ; Search time 4821.79 Seconds
(without alignments)
6494.048 Million cell updates/sec

Title: US-09-835-082-1
Perfect score: 2320

Sequence: 1 ctcagctgcagttctctgat.....ttctgatgtttctctacaact 2320

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 13736207 seqs, 6748477542 residues

Total number of hits satisfying chosen parameters: 27472414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_estl:*
10: gb_estt:*
11: gb_htc:*
12: gb_gss:*
13: em_gss_hum:*
14: em_gss_inv:*
15: em_gss_pln:*
16: em_gss_vrt:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1631.6	70.3	2450	11	BC011409 Homo sapi
2	789.6	34.0	2306	11	BC012716 Mus muscu
3	780.2	33.6	2202	11	AK002629 Mus muscu
4	741.2	31.9	868	10	BC402908
5	722.4	31.1	882	10	BC291462
6	720.8	31.1	876	10	BI759297
7	707.4	30.5	884	10	BI762347
8	695.8	30.0	818	9	AI307269
9	695.4	30.0	881	10	BC289777
10	680.6	29.3	1005	10	BE870911
11	679.2	28.3	697	10	BE747143
12	669.4	28.9	764	9	AI831562
13	644	27.8	723	10	BC424452
14	641.2	27.6	829	10	BC469506
15	640	27.6	833	10	BC400591
16	639.8	27.6	791	9	AI765716
17	636.8	27.4	658	9	AV652040

18	635.4	27.4	945	10	BC167885
19	625.6	27.0	724	10	BF031444
20	621.8	26.8	688	10	BC286763
21	618.8	26.7	796	10	BC432363
22	604.8	26.1	861	9	AI829268
23	601	25.9	647	9	AV692287
24	596	25.7	763	10	BC399019
25	594.2	25.6	663	9	AV646537
26	592.4	25.5	859	10	BE249993
27	589.8	25.4	640	9	AI635995
28	588	25.3	683	9	AI480360
29	584.8	25.2	812	10	BC401258
30	584	25.2	621	9	AW151709
31	583	25.1	667	9	AW051715
32	583	25.1	969	10	BC291839
33	572	24.7	636	9	AI767470
34	567.4	24.5	581	10	BE379923
35	567	24.4	623	10	BC171502
36	557.8	24.0	743	10	BC386668
37	556.4	24.0	741	10	BC433491
38	550.6	23.7	619	9	AW300312
39	549.8	23.7	851	10	BC401340
40	543.8	23.4	700	10	BF130433
41	542.8	23.4	608	9	AW469219
42	542.8	23.4	953	10	BC328092
43	537	23.1	935	10	BF234991
44	535.4	23.1	626	9	AW082933
45	532.2	22.9	666	10	BF027719

ALIGNMENTS

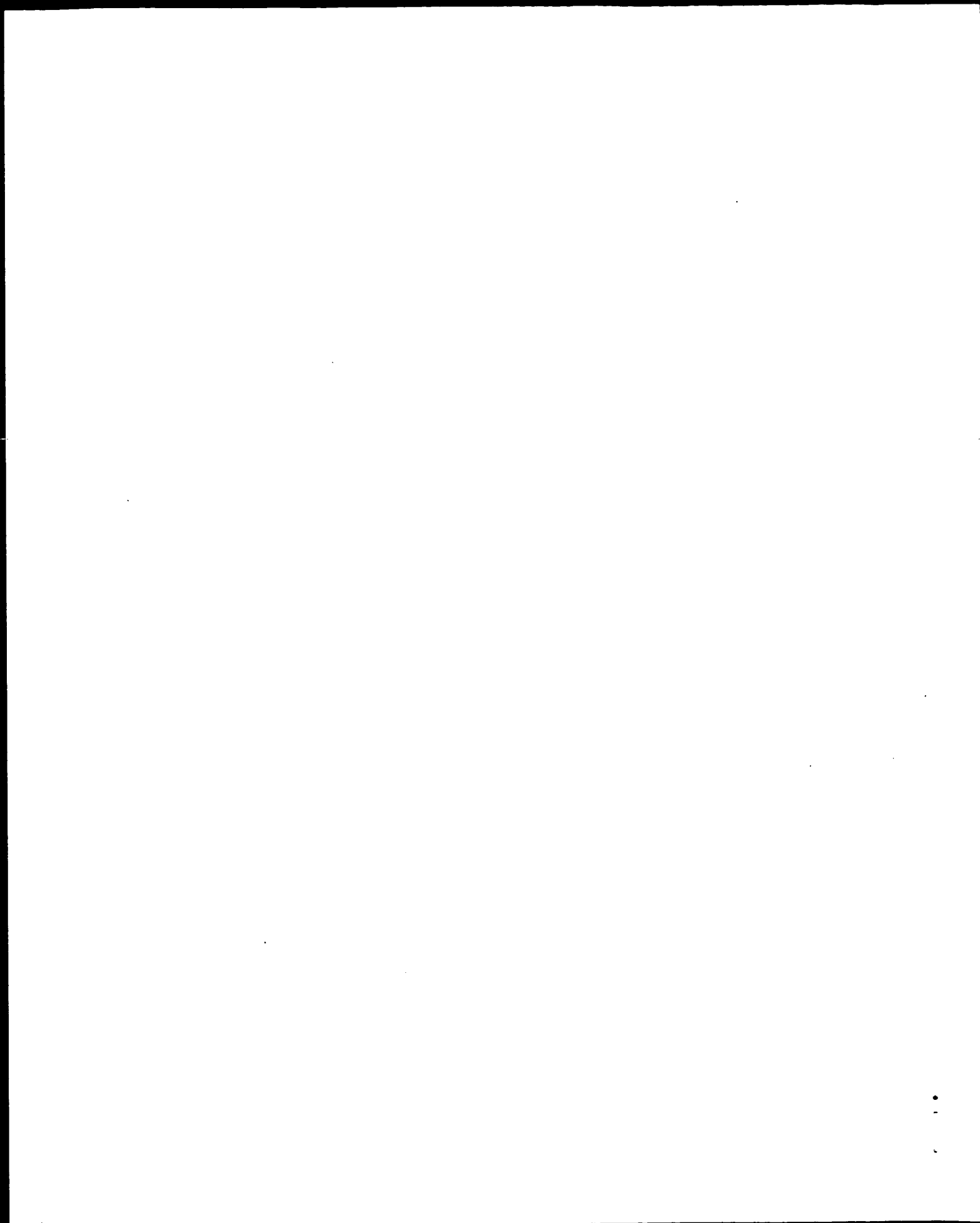
RESULT 1

BC011409
LOCUS Homo sapiens, clone IMAGE:3854370, mRNA. linear HTC 30-JUL-2001
DEFINITION BC011409
ACCESSION BC011409.1 GI:15030273
VERSION HTC.
KEYWORDS human.
SOURCE
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
Strausberg, R.
Direct Submission
Submitted (23-JUL-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA
NIH-MGC Project URL: <http://mgc.nci.nih.gov>
Contact: MGC help desk
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)
Sequencing Center
Center code: BCM-HGSC
Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>
Contact: villalon@bcm.tmc.edu
Villalon, D.K., Luna, R.A., Hale, S.M., Hulyk, S., Lu, X., Garcia, A.M., Holloway, M., Telford, B., Hodgson, A., Bouck, J., Yu, W., Muzny, D.M., Gibbs, R.A.

REMARK COMMENT

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Series: IRAC Plate: 13 Row: b Column: 4
This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 4507814
This clone has the following problem: incomplete processing.
Location/Qualifiers

FEATURES



1459	taccagtaacattcctcttgacatgatgtttctctcttgccgctgctgctgacatggcc	1518
QY		
1460	ttct	1519
DB		
1441	TACCAGTACCAATCCTTTGACGTATGGTGTCTCTTTGGCCGCTCGTGTGACATGGCC	1500
DB		
1519	ttcatacctttaaatgtgtcttagctaccggaatctctggggaaaaagggcga	1578
QY		
1501	TTCAATCACCTTTAAATGTTGTCTTATGGTACCGGAAATGCTTTGGGAAAAAGGGCGA	1560
DB		
1579	gttaagaagcccaaaatccaagaccattga	1611
QY		
1561	GTTTAAGAAGCCCAAAATCCAAGACCCATTGA	1593
DB		

Search completed: August 21, 2002, 11:58:40
Job time: 32903 sec

Db 299 AATACTACACAGAAAATCTATCCAGTGCCTGATGACCAAGAGAGTGGAAGAACCGTT 358
Qy 287 tcaaggcttttgcacatctcaatggaagcacaagtagaagtagatatattctctattaa 346
Db 359 ACCAATCAATTTGGAACAATCACTTGTGTAGCGATCATTCCTAACTGCTCTCAGACAG 418
Qy 347 tgggttcatacaatgaca---tttttgacttatttttcaaatggaagattgtttta 403
Db 419 AGTACAGGAAATACATGATTGTTATTTGGCTGTACTTCATCAACTGCCAGAGCCTCTGC 478
Qy 404 aagacaaaaattagtagaataacttaaggagagttctttttgatcagtgtttctogac 463
Db 479 AGGACAGGACACCCCTGAATCTTTTAAGGAGAGCAAGTTGATGCTCTTTTCACAGAC 538
Qy 464 cttttgataactgtgcttaattgttgcaaatatttccctccctccctccctccctccct 523
Db 539 CAGCCTTACCTGTGGGTGATCTCGCTGAGTATTTGGGCTTACCATCTGTGTACTCT 598
Qy 524 ccagggaatactttgacactatcttgaaagagtgacagtgccctgctctctcttct 583
Db 599 TCAGGGGTTTCCGTGTCCTGGAGCATACATTCAGCAGAAAGCCAGACCTGTGTCT 658
Qy 584 atgtcccagaattctctaggttctcagatgccatgacttttaaggagagagtagcaga 643
Db 659 ACATTCAGGTGTACACAAAGTTTCAGACACACATGATCTTTCCACAGAGTGGCCA 718
Qy 644 accaatcatgacttgaggagaaactttattatgcccacgcttttttcaaaaaatgcccag 703
Db 719 ACTTCCTGTTAAATTTTGGAGCCCTATCTATTTATTTGCTGTTTCAAAAGTATGAGA 778
Qy 704 aaatagctctgaaattctccaaacacactgttaaggagtagatctctacagccacat 763
Db 779 AACTGCATCAGCTGCTCAGACAGAGATGT---GGATATATCACTATTCAGAGGTCT 835
Qy 764 caatttggttggaacgagactttgtttggaactatcccaaacccogtggatcccaaca 823
Db 836 CTGTTTGGCTGTTAAGATATGACTTTGTGCTGAATATCTTAGGCGGCTATGCCAACA 895
Qy 824 tgatctcattggttggtatcaactgccatcagggaaagcgttgctctatggaattggaag 883
Db 896 TGTCTTCATTTGGAGGTATCACTGTAAGAAGAGGAAGAACTTGTCTCAGGAATTTGAAG 955
Qy 884 cttacatgaattgtcttggaacacatggaattggttttctctcttggtggaatgaattgct 943
Db 956 CTTACATTAATGCTTCTGAGAACATGGAATGTGGTGTCTTCTTTGGGATCAATGGTCT 1015
Qy 944 cagaattccagagaagaagctatggcaattgctgactgcttggtggcaaaatccctcaga 1003
Db 1016 CAGAAATCCAGAGAGAAAGCTATGGCAATTGCTGATGCTTTGGCAAAACCCCTCAGA 1075
Qy 1004 cagtctgtggcgttacactggaacccgaccatcgaaatcttgcaacaacacagatacttg 1063
Db 1076 CAGTCTGTGGCGGTACACTGGAACCCGACCATCGAATCTTGGCAACACACAGATACTTG 1135
Qy 1064 ttaagtgtctaccacaaacgatctgcttggtcacccgatgacccgtgcttttaccacc 1123
Db 1136 TTAAGTGTCTACCCAAAACGATCTGTTGTCACCCGATGACCCGCTTATCACCC 1195
Qy 1124 atgctggttcccatggtgtttatgaaagcatatgcaatgggttcccatggtgatgac 1183
Db 1196 ATGCTGTGTTCCATGTGTTTATGAAGCATATGCAATGGCTTCCCATGCTGATGATGC 1255
Qy 1184 cttgtttggtgacagatggacaatgcaagcgcatggagactaaaggagctggagtga 1243
Db 1256 CTTGTTTGGTGTACAGATGGACATGCAAGCCCATGGAGACTAAGGGAGCTGGAGTGA 1315
Qy 1244 ccctgaattgtctggaattgaactctcgaagatttagaaaatgctcttaaaagcagtcata 1303
Db 1316 CCTGAATGTCTGGAATAGACTCTGTAAGATTAGAAAATGCTCTAAAAGCAGTCATCA 1375
Qy 1304 atgacaaaagttaacaggaggaacatcatgagcctctccagccttcaaaaggaaccccg 1363
Db 1376 ATGACAAAAGTTACAGGAGAACATCATGCGCTCTCCAGGCTCTCCAGGCTCACAAGCCGCCG 1435

Qy 1364 tggagcccttgacctggcgtgttctggttgaggtttgtgatgaggcacaagggcgccg 1423
Db 1436 TSGAGCCCTGACCTTGCCGCTGTTCTGGGTGGAGTTTGTGATGAGGCACAAAGGGCGCC 1495
Qy 1424 cacactcgcccccgagccacgacctcactcgtgtaccagttaccattccttggagctga 1483
Db 1496 CACACTCGCCCGCCGACGCCACGACCTCACTGTGTACCACTACCATTCCTTGGAGCTGA 1555
Qy 1484 ttggttctcttggcgtgtgctgacagtggtccttcacacacacacacacacacacacac 1543
Db 1556 TTGTTTCTCTTGGCGCTGCTGCTGACAGTGGCTTCATCACTTTAAATGTTGCTCTT 1615
Qy 1544 atggctacccggaatgcttgggaaaaaagggcgagtgaaagaaagccacacacacacac 1603
Db 1616 ATGGGTACCCGAAATGCTTGGGAAAAAAGGGGAGTAAAGAAAGCCACAAATCCCAAGA 1675
Qy 1604 cccattgagaagtggtgggaaataaggtaaaaattttgaacattccctagtcatttcca 1663
Db 1676 CCCATTGAGAGTGGTGGGAAATAAGGTAAATTTTGAACCA-TCCCTAGTCATTTCCA 1734
Qy 1664 aacttgaaaaacagaaatcagtgtaaatcattttatttatttatttatttatttatttattt 1723
Db 1735 AACTTTGAAAAACAGAAATCACTGTTAAATTCATTTTATTTATTTAAGGAAATACTTTGCAT 1794
Qy 1724 aaatttaatacagccccagagtg-ctttaaaaaaattctttaaataaaaaataatagactgc 1782
Db 1795 AAATTAATCAGCCCCAGAGTGTCTTTAAAAAATTCCTTTAAATAAAAAATAATAGACTGC 1854
Qy 1783 tagtcagtaaaagatatatttaataatgatatcgtgccctcccggtctcttggatcagatga 1842
Db 1855 TAGTCAGTAAAGATATTTGAATATGATATGTCGTCGCCCTCCGGAGTCTTTGATCAGATGA 1914
Qy 1843 catgtgccatttttcagagagcgtgcagacagcgtgcattctagatttacttttctact 1902
Db 1915 CATGTGCCATTTTCAGAGAGAGCTGCAGACAGCTGGCATTTAGATTTACTTTTCTTACT 1974
Qy 1903 ctgaaacatgacctgtttggagtgctgggattcaaaagtggtccacacgctgcccact 1962
Db 1975 CTGAAACATGCGCTGTTTGGGAGTGGGGATTCAAAGGTGGTCCACGCTGCCCTACT 2034
Qy 1963 gcaatggcagtttaattcttattcttctgttctcgtcagatggttgcaattgactccttaa 2022
Db 2035 GCAAAATGCGAGTTTAAATCTTATCTTTTGGCTTCTGCAGATGGTTGC-ATTGATCCTTAA 2093
Qy 2023 ccaataatggttcagctcctcctctctctg-cttcattagtgccacctgtgtgttttaa 2081
Db 2094 CCAATAATGTCAGTCTCTCATCTCTGCTGACTTCTATAGGTGGCCACCTTGTG-GTTTAA 2152
Qy 2082 agaaggaagccttgtacacttttagagtgtagtgaaatgaatgaatgagtgagtgcaac 2141
Db 2153 AGAAGGGAAGCTTTGTACCTTTAGAGTGTAGGTGAAATGAATGAATGGCTTGGAGTGCAC 2212
Qy 2142 tgaagaacagcatatgattcttctgttgggaaaaaagaatgactgataaattggtggg 2201
Db 2213 TGAGAACAGCATATGATTTCTTGTGTTTGGGAAAAAAGAAATGCTATGAATTTGGTGGG 2272
Qy 2202 tgggtatttgagaagataatcattgcttattgttcaaatggagctgaatttgataaaaaacc 2261
Db 2273 TGGTGTATTTGAGAAGATAATCATTTGCTTATGTCAAAATGGAGCTGAATTTGATAAAAAACC 2332
Qy 2262 caaatacagctatgaagtgtgtgggcaagtttac-ttttttctgagtttctcctacaac 2320
Db 2333 CAAAATACAGTATGAAGTGTCTGGGCAAGTTTACTTTTTTCTTGATGTTTCTCTACACT 2392

RESULT 15
AF030310

LOCUS

AF030310 Homo sapiens UDP-glucuronosyltransferase 1A8 (UGT1A8) mRNA,
complete cds.

ACCESSION

AF030310.1 GI:2613043

Db 1256 CCTTGTGTTGATCAGATGGCAATGCAAGCCGATGGAGACTAAGGAGCTGGAGTGA 1315
 QY 1244 cccagaagtcttggaataactctgaagatttagaataatgctctctaaagcagtcata 1303
 Db 1316 CCTGGAATGTTCTGGAATGACTTCTGAGATTTAGAAATGCTCTAAAGCAGTCATCA 1375
 QY 1304 atgacaaaagttaacaaggaagacatcgccctctccagctctcaagaccgcccgg 1363
 Db 1376 ATGACAAAAGTTACAAGGAGAACATCGCCCTCTCCAGCCCTTCAAGGACGCCCGG 1435
 QY 1364 tggcgccctggaactgcccgtgtctgtggttgagtttggatgagggcacagggcgcc 1423
 Db 1436 TGGAGCCGCTGGACTCGCCGTGCTGCGGTGGAGTTGTGATGAGGCACAAAGGCGCG 1495
 QY 1424 cacactcgccgcccagccagccactcactgtgtaccagttaccattcttgaagtgga 1483
 Db 1496 CACACCTGGCCCGCCAGCCAGCCCTCAGCTGTACACAGTACCATCTCTTGAGCTGA 1555
 QY 1484 ttggttctcttgcccgctgctgacagtggtggtggtggtggtggtggtggtggtggt 1543
 Db 1556 TTGTTTCTCTTGGCCGTGCTGCTGACAGTGGCCTTCATCACCTTTAAATGTTGCTTT 1615
 QY 1544 atggctacgggaatgcttggaagaaagggcgagttgaagaaagcccaaatccaaga 1603
 Db 1616 ATGGCTACCCGAATGCTTGGGAAAGAGGCGAGTTAGAAAGCCCRCAATTCACAGA 1675
 QY 1604 cccattgagaagtggttggaataaaggtgaaattttgaacattccctgattcttca 1663
 Db 1676 CCATTGAGAGGTGGTGGGAAATAGGTAAATTTTGAACCA-TCCCTAGTCAATTCGA 1734
 QY 1664 aacttgaagaaacagtaactggttaaatcatttcttcttatttgaagaaacttggat 1723
 Db 1735 AACTTGAAGAACAGATCAGTGTAAATTCATTTATTCATTATAGGAAATACCTTTGCA 1794
 QY 1724 aaatttaacagccccagtg-ctttaaaaaattctcttaataaaaaataagactcgc 1782
 Db 1795 AAATTAATCAGCCAGAGTGCTTTTAAATAATCTCTTAAATAAATAATAGACTCGC 1854
 QY 1783 tagtcagtaagataatttgaatgatctatcgccctccgctgcttcttgaagata 1842
 Db 1855 TAGTCAGTAAAGATATTGAATATGATGTCGTCGCCCTCCGAGTCTTTGATCAGATGA 1914
 QY 1843 catgctcatttttcagagagcgtgcagacagcgtggtggtggtggtggtggtggtggt 1902
 Db 1915 CATGTCGCATTTTTCAGAGAGCGTGACAGAGCGTGGCATTCTAGATTCTTTCTTACT 1974
 QY 1903 ctgaaacacatgctgtgttggtgagtggtggtggtggtggtggtggtggtggtggt 1962
 Db 1975 CTGAAACATGCGCTGTTTGGGAGTGGGGATTCAAGAGTGTGCCACGGCTGCCCTACT 2034
 QY 1963 gcaaatgacagtttgaatcttcttcttcttcttcttcttcttcttcttcttcttct 2022
 Db 2035 GCAATGCGAGTTTAACTTAACTTTTGGCTTCTGAGATGGTTC-ATTGATCTCTAA 2093
 QY 2023 ccaataatgctcagctcctcctctgctgctgctgctgctgctgctgctgctgctgct 2081
 Db 2094 CCAATATGCTAGTCTCTCATCTGCTCTGCTGCTTCAATAGTGGCAGCTTTGTG-GTTTAA 2152
 QY 2082 agaaaggaagctttgaccttagagtgtagtgtagtgtagtgtagtgtagtgtagtgtag 2141
 Db 2153 AGAAGGAGCTTTGACCTTACCTTACCTTACCTTACCTTACCTTACCTTACCTTACCT 2212
 QY 2142 tgaagacagatgatattcttctgcttctggtggtggtggtggtggtggtggtggtggt 2201
 Db 2213 TGAGAACAGGATATGATTTCTTGTGTTGGGAAAGAAAGATGCTATGAATTTGGTGGG 2272
 QY 2202 tggttatttgaagataaactcttcttcttcttcttcttcttcttcttcttcttcttct 2261
 Db 2273 TGGTGTATTGAGAGATAATCATCTGTTATGCAAAATGGAGCTGAATTTGATAAAACC 2332
 QY 2262 caaantacagctatgaagtgctggggaagtttac-ttttttctgagtttcttctcaact 2320
 Db 2333 CAAATACAGCTATGAAGTGTGGGCAAGTTTACTTTTTTCTGATGTTTCTTCAACT 2392

RESULT 14

HUMUGTIFA

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

COMMENT

FEATURES

source

gene

CDS

HUMUGTIFA 2422 bp mRNA linear PRI 14-JAN-1995
 Homo sapiens phenol UDP-glucuronosyltransferase (UDPGT) mRNA,
 complete cds.
 J04093
 J04093.1 GI:624724
 UDP-glucuronosyltransferase; phenol.
 Homo sapiens kidney cDNA to mRNA.
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 2422)
 Harding, D., Fournel-Gigleux, S., Jackson, M.R. and Burchell, B.
 Cloning and substrate specificity of a human phenol
 UDP-glucuronosyltransferase expressed in COS-7 cells
 Proc. Natl. Acad. Sci. U.S.A. 85 (22), 8381-8385 (1988)
 89042187
 On Jan 18, 1995 this sequence version replaced gi:340517.
 Location/Qualifiers
 1..2422
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /tissue_type="kidney"
 1..2422
 /gene="UDPGT"
 88..1683
 /gene="UDPGT"
 /EC_number="2.4.1.17"
 /codon_start=1
 /product="UDP-glucuronosyltransferase"
 /protein_id="AAA61251.1"
 /db_xref="GI:624725"
 /translation="MACLLRSQRIASGVFFALWGVYGVKLLVVPQDSHWLSMKD
 IVEVLSDRGHEIVVYVEVPEVLLKEYKYYTRKIPYDQELKNRYOSFGNNHFAER
 SELTAPQTEYRNNMIVIGLTFINCSLLQDRDTLNFKEKEDALFTDPALPCGVILA
 EYGLSVYLFGRFPFCSLEHTSRSDPVSYIPRCYTKFSDHMTFSRVANFLVNLLE
 PYLFCLFSEKYLASAVLRDVIDITLSEVSWLLRYDFVLEYPYPMVMVPIGIGI
 NCKRKDLISQEFAYINASGHEIVVFSIGSMVSEIPEKKMAIADALGKNPQTVLWR
 YTGTRPSNLANTILVKWLPONDLLGHPTAFITHAGSHGVYESICNGVPMVMPPLF
 GQMDNAKRMETGAGVTNLVLEMTSEDLNALKAVINDKSYKENIMRLSLKHDRPV
 EPDLIAVFWFVYMRHKGAPHLRPAADLTWYQVHSLDIVGLLAVLVLTAFITFKCC
 PVGYPKCLGKGGKVRKAKHKSHT"
 2396..2401
 /gene="UDPGT"
 2404..2409
 /gene="UDPGT"
 2422
 /gene="UDPGT"
 653 a 516 c 555 g 698 t

polyA_signal

polyA_signal

polyA_site

BASE COUNT

ORIGIN

Query Match 66.3%; Score 1539.2; DB 9; Length 2422;
 Best Local Similarity 82.0%; Pred. No. 0;
 Matches 1870; Conservative 0; Mismatches 398; Indels 12; Gaps 8;

QY 47 ttccctctatgtgtgtctgctgctgacacctgtgcttgcagagggcagggagactactgg 106

Db 119 TTTCTGCAGGGGTTTCTTCTTAGCTCTTTGGGCGATGTTGTAGTGACAGCTGCTGG 178

QY 107 tagtgcacagatggagaccacgtggttaccatgaggtcggtggtggagaaactattc 166

Db 179 TGGTCCCTCAGGACGAGGACCCACCTGGCTTAGTATGAAGATATAGTTGAGGTTCTCAGTG 238

QY 167 tcaaggggcagaggtggtgttagtcatcgacagaggtgagttggcaactgggaagatcac 226

Db 239 ACCGGGGTCATGAGATTGTTAGTGTGTGCTGCTGAAGTTAATTTCTTTTGAAGATACA 298

QY 227 tgaattgcacagtggaagacttatccaacttcataaccctggaggtcctggaccggaggt 286


```

Db 726 AACTCCGATCAGCTGCTCCTCAAGAGAGATGTGGATATATATCAACCTTATATCAGAGGTCT 785
Qy 764 caatttggttggtgcgaacggaacttggtttgactatcccaaacccggtgatgcccaaca 823
Db 786 CTGTTTGGCTGTTAGATATGACTTGTGCTTGAATATCTAGCCGGTCAATGCCCAACA 845
Qy 824 tgatcttcatttggttgatatacaactgcacatcgaggaaagccggttgacctatggaattgaag 883
Db 846 TGGTCTTCATTTGGAGGTATCAACTGTAAAGAGAGAAAGACTTGTCTCAGGAATTTGAAG 905
Qy 884 cctacataattactctctggagaacatggaattggtttctctctctctctctctctctctctct 943
Db 906 CTTACATATATCTCTTGGAGACATGGAATGTGGTTTCTCTTGGGATCAATGTGCT 965
Qy 944 cagaaatccagagaagaagctatggaattgctgatgctttgggcaaaatccctcaga 1003
Db 966 CAGAAATCCAGAGAAGAAAGCTATGGCAACTGTGTATGTCTTGGGCAAAATCCCTCAGA 1025
Qy 1004 cagtctctgtggcggtacactggaaacccgaccatgaatctctggaacacacgatacttg 1063
Db 1026 CAGTCTGTGGCGGTACACTTGGAAACCCGACCATCGAATCTTGGCAACACGATACTG 1085
Qy 1064 ttaagtggtaaccccaaaacgatatctgttgtaacccgatagcacctgacctttatccccc 1123
Db 1086 TTAAGTGGCTACCCCAAAACGATCTGCTTGGTCAACCGGATGACCGTGGCTTTATCACCC 1145
Qy 1124 atgctggttcccatggtttatgaaacatataatgagctgcttccctggtgatgacgc 1183
Db 1146 ATGCTGTTCCCATCGGTGTTTGAAGGCATATGCATGGCTGCCCTCCATGGTGTATGTC 1205
Qy 1184 cctgttggtgatcagatggacaatgcaagcgcatggagactaaggagctgagtgga 1243
Db 1206 CTTGTTTGGTGATCAGATGACAAATGCAAGCGCATGGAGACTAAGGAGCTGGAGTGA 1265
Qy 1244 cctgtaattctgaaatgactctgaaatttagaaatctctaaagcagtcac 1303
Db 1266 CCCTGATGTTCTGGAAATGACTCTGAAGATTTAGAAATGCTCTAAAGCAGTCATCA 1325
Qy 1304 atgacaagaattacaaggagaacatcatgcgctctccagctctccacaaaggacgcgcgg 1363
Db 1326 ATGACAAAAGTTACAAAGAGAACATCATGCGCTCTCCAGCCTTCAAGAGACCGCCCGG 1385
Qy 1364 tggagcgcgtgacacgtcgctgtctggtggtgagttgtgatggagcaagggcgcg 1423
Db 1386 TGGAGCGCTGGACCTGCGCTGTTCTGGGTGGAGTTGTGATGAGGCACAAAGGCGCGC 1445
Qy 1424 cacaactgccccccagccacagacactcactggtaccagttaccattccttgacatga 1483
Db 1446 CACACCTGCGCCCGCCAGCCACCCACGACTCACCTGTGTACAGTACCATTCCTTGGACGTGA 1505
Qy 1484 ttggttctcttggcgctgctgacagtggtgcttccatccactttaaattgtgtgctt 1543
Db 1506 TTGTTTCTCTTTGGCGCTGCTGCTGACAGTGGCTTCAACCTTTAAATTTGTTGCTT 1565
Qy 1544 atggtacagaaatgcttgggaaagagcaggttaagaaagcccaaaatccaaga 1603
Db 1566 ATGCTACCGGAAATGCTTGGGAAAGAGGCGAGTTAAGAAAGCCCAATCCAAAGA 1625
Qy 1604 cccattgagaagtgggtgggaaataaggttaaaatttgaaccattcccttagctattcca 1663
Db 1626 CCCATTGAGAAGTGGTGGGAAATAAGTAAATTTGAACCATTCCTTAGTCATTTCGA 1685
Qy 1664 aacttgaaacagaatacgttgtaaaattcattttatttcttataagaaataactttgat 1723
Db 1686 AACTTGAACACAGATCAGTGTAAATCATTTTATTCTTATTAAAGAAATTAATTTGCAT 1745
Qy 1724 aaattaatcagccccagagtgctttaaataattctcttaaataaaaaataatagactgcgt 1783
Db 1746 AAATTAATCAGCCCGAGTCTTTAAAAATTTCTTTAAATAAAAAATAATAGACTCGCT 1805
Qy 1784 agtcagtaagattgaattatgatctgccccctcagtgctcttgatcagatgac 1843

```

```

Db 1806 AGTCACTAAAGATATTTGAATATGTATCTGTCGCCCTCCGCTGCTTTGATCAGATGAC 1865
Qy 1844 atgtgccatttttcagaggagcgtgcagacagcgttgcgaccttagattacttttcttact 1903
Db 1866 ATGTGCCATTTTTCAGAGGACGTCGAGACAGGCTGCAATCTAGATTACTTTCTTACTC 1925
Qy 1904 tgaacaatgaccttcttgggagtgcggaattcaaaagtggtccaccgctccctactg 1963
Db 1926 TGAACATGGCGCTGTTGGAGTGGCGGATTCAAAGGTGTCCACCGCTGCCCTACTG 1985
Qy 1964 caatggcagttttaaacttctattcttggctctgcagatgggttgcaattgatcttaac 2023
Db 1986 CAAATGGCAATTTTAAATCTTATCTTTGGCTCTGCAATGGTTCGAATGTATCTTAAAC 2045
Qy 2024 caataatggtcagtcctcatctctctctctctctctctctctctctctctctctctct 2083
Db 2046 CAATAATGTCAGTCCCTCATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2105
Qy 2084 aagggaagcttctacaccttagagtgtagtggaatgaatgaatgaatgaatgaatgaatga 2143
Db 2106 AAGGGAAGCTTTTACCTTTAGAGTGTAGGTGAAATGAATGAATGAATGAATGAATGA 2165
Qy 2144 agaacacatgatgttctctctctctctctctctctctctctctctctctctctctctct 2203
Db 2166 AGAACAGCATATGATTTCTTCTTGGGAAAGAAATGATGCTATGAAATGGTGGTG 2225
Qy 2204 gtgtattgagaataatcattgcttctctctctctctctctctctctctctctctctct 2263
Db 2226 GTGTATTGAGAAGATAATCATTGCTTATGTCAAATGGAGCTGAATTTGATAAAACCCA 2285
Qy 2264 aatacagctatgaatgctgggcaagtttactttctctctctctctctctctctctctct 2320
Db 2286 AAATCAGCATGATGAAGTGTGGGCAAGTTTACTTTTCTGATGTTTCTTCTTCTTCT 2342

```

RESULT 12

```

HUMHUGBR1
LOCUS Human bilirubin UDP-glucuronosyltransferase isozyne 1 mRNA,
DEFINITION complete cds.
ACCESSION M57899.1 GI:18472
VERSION UDP-glucuronosyltransferase; bilirubin UDP-glucuronosyltransferase.
KEYWORDS Human adult female liver, cDNA to mRNA, clones 26, 211 and 26MB2.
SOURCE Homo sapiens
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 2351)
AUTHORS Rittner,J.K., Crawford,J.M. and Owens,I.S.
TITLE Cloning of two human liver bilirubin UDP-glucuronosyltransferase
CDNAS with expression in COS-1 cells
JOURNAL J. Biol. Chem. 266 (2), 1043-1047 (1991)
MEDLINE 91093210
FEATURES
Location/Qualifiers
1..2351
/organism="Homo sapiens"
/isolate="AK"
/db_xref="taxon:9606"
/map="2"
/clone="26"
/sex="female"
/tissue_type="liver"
/dev_stage="adult"
/tissue_lib="lambda-ZAP"
1..2351
/gene="UGT1"
16..1617
/gene="UGT1"
/standard_name="bilirubin UDP-glucuronosyltransferase
isozyme 1"
/EC_number="2.4.1.17"
/codon_start=1
/product="UDP-glucuronosyltransferase 1"

```

```

Db 1901 GGCATTCTAGATTACTTTCTTACTCTCAACATGCGCTGTTGGAGTGGGGATTCAA 1960
Qy 1938 aggtggtccaccgctgcccctactgcaaatggcaggttttaattcttcttcttggctct 1997
Db 1961 AGGTGTCCTCCACCGCTGCCCTACTGCAATGGCAGTTTAACTTATCTTTGGCTTCT 2020
Qy 1998 gcaagtgttgcattgaccttaaccaataatggctcagctcctctctctctctctc 2057
Db 2021 SCAGATGGTTCGCAATGATCCTTAACCAATATGGTCAGTCTCATCTGCTCTCTTC 2080
Qy 2058 atagtgccaccttgggttttaagaagggaagcttttgcacctttagagtgtagtgaa 2117
Db 2081 ATAGTGGCCACCTTGTGTGTTTAAAGAAAGGAAGCTTTGTACCTTTAGAGTGTAGTAA 2140
Qy 2118 atgaatgaatgcttgagtgactgagacagacacatattctctgcttcttgggaaaaa 2177
Db 2141 ATGAATGAATGGCTTGGAGTGACACTGAGACACAGCATATGATTTCTGTTGGGAAAAA 2200
Qy 2178 gaatgatctatgaattgggtgggtgggtgatttattgagaagataatcattgcttatgaa 2237
Db 2201 GAATGATGCTATGAATTTGGTGGGTGGTGTATTTTGAAGAAGATAATCATTTGCTTATGCAA 2260
Qy 2238 atggagctgaatttgataaaaccccaatacacagctatgaagtctgggcaagtttactt 2297
Db 2261 ATGGAGCTGAATTTGATATAAAACCCCAAAATACAGCTATGAAGTGTGGCAAGTTTACTT 2320
Qy 2298 tttttctgatgtttctcaact 2320
Db 2321 TTTTCTGATGTTTCTTACAAC 2343

RESULT 11
BC019861 2388 bp mRNA linear PRI 22-JAN-2002
LOCUS Homo sapiens, clone MGC:29860 IMAGE:4454322, mRNA, complete cds.
DEFINITION BC019861
ACCESSION BC019861
VERSION BC019861.1 GI:18043700
KEYWORDS human.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 2388)
Strausberg, R.
Direct Submission
Submitted (19-DEC-2001) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NIH-MGC Project URL: http://mgc.nci.nih.gov
Contact: MGC help desk
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Baylor College of Medicine Human Genome
Sequencing Center
Center code: BCM-HGSC
Web site: http://www.hgsc.bcm.tmc.edu/cdna/
Contact: amg@bcm.tmc.edu
Guaratine, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Hale, S.M.,
Yoon, V.S., Kowis, C.R., Lawrence, S., Martin, R.G., Muzny, D.M.,
Richards, S., Gibbs, R.A.

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Series: IRAC Plate: 42 Row: c Column: 8
This clone was selected for full length sequencing because it
passed the following selection criteria: matched mRNA gi: 4507814.
Location/Qualifiers
1. .2388
/organism="Homo sapiens"
/db_xref="taxon:9606"

```

```

/clone="MGC:29860 IMAGE:4454322"
/tissue_type="kidney, hypernephroma"
/clone_lib="NIH_MGC_89"
/lab_host="DH10B"
/note="Vector: pCMV-SPORT6"
35..1633
/codon_start=1
/product="unknown (protein for MGC:29860)"
/protein_id="AAH19861.1"
/db_xref="GI:18043701"
/translation="MACLLRSFORISAGVFFLALMGWVMDKLLVVPQDGGSHWLSMKD
IVEVLSDRGHEIVVYVPEVNLILKESKYTRKIYVYDQELKNRYOSFGNNHFAER
SEVTAQDHEIVNNMIVIGLYFINCSLLQDRDLTFKESKDFALTPALPCGVILA
EYLGLPSVLFGRFPCPSLEHTFSRSPDPVSIPRCYTKSDHMTFSQVAVNLFVNLLE
PLFYCLFSKYVELASAVLKROVDIITLYQKVSWLLRDFVLEYPRVPMNVFVIGG
INCKRKDLDSQEPAYINASGHBGIVFESLGSWSEIPEKKAMATADALGHIPQVULM
RYTGRPNLANNTILYKWLPNQDILLGHPTAFITIHAGSHGYIESICNGVPVYMPML
FGDDMDNAKRMETKGAVTLNVLMTSESDLENALKAVINDKSTYKENIMRLSSLKHDP
VEPLDLAVFWFEVFEVVRHKGAPHLRPAAHDLTWQYHSLDVIGFLAVVLTVAFTFKC
CAYGYRKLGLKGRVKKAKHKSHT"
BASE COUNT 648 a 503 c 546 g 691 t
ORIGIN

Query Match 70.2%; Score 1629; DB 9; Length 2388;
Best Local Similarity 82.5%; Pred. No. 0;
Matches 1879; Conservative 0; Mismatches 395; Indels 3; Gaps 1;

Qy 47 ttctctatgtgtgtctgtcgtgaacctgtgcttccagaggcagggaagctactgg 106
Db 66 TTTCTGCACGGGTTTCTTCTTAGCACCTTTGGGGCATGGTTGTAGGTGACACGCTCTGG 125

Qy 107 tagtcccatgatgaggaacactggttcaccatgaggtcggtggtgagaaactcatic 166
Db 126 TGTCCCTCAGAGGAGGACGACTGGCTTAGTAGAAGGATATAGTGTGCTCTCAGTG 185

Qy 167 tcagggggcatgaggtggtgttagtcagcagaggtgagttggcaactgggaagatcac 226
Db 186 ACCGGGGTCATGAGATTGTAGTGGTGGTGGCTGAAGTTAATTTGCTTTGAAAGAAATCCA 245

Qy 227 tgaattgcacagtgaagacttattcaacttcataccctgagagatctgagccggagt 286
Db 246 AATACTACACAGAAATAATCTATCCAGTGGCGGTATGCCAAGAAAGCTGAGNACCGTT 305

Qy 287 tcaaggtcttgcacatgctcgaatggaagcacagtacagtaagtataattctctattaa 346
Db 306 ACCAATCATTTTGGAAACATCACTTTGCTGAGCGATCATTTCTTAACCTCTCTCAGACAG 365

Qy 347 tgggttcatacaatgaca---ttttgacttatttttttcaaatgagaggtttgttta 403
Db 366 AGTACAGGATTAACATGATTGTTATTGGCTGTACTTCACTCAACTGCCAGAGCTCTCTGC 425

Qy 404 agacacaaattagtagaataacttaaggagaggtttttttagtcagtggtttctcgatc 463
Db 426 AGGACAGGACACCCCTGAACATTTCTTTAAGGAGAGCAAGTTTGTATGCTCTTTTCACAGACC 485

Qy 464 cttttgataactgtggcttaattgttgccaaataatttctccctccctccgtggtcttg 523
Db 486 CAGCCTTACCTGTGGGGTGATCCTGGCTGAGTATTTGGGCTTACCATCTGTGTACTCT 545

Qy 524 ccaggggaatactttgcccactatcttgaaagggtgcacagtgccctgctctctcttct 583
Db 546 TCAGGGGGTTTTCGGTGTTCCTTGGAGCATACATTTACAGAGAGAGCCCGCTGTGCTCT 605

Qy 584 atgtcccccgaattctcttaggtttctcagatgccatgcttctcaaggagagatcagga 643
Db 606 ACATTTCCCGAGGTGTACACAAAGTTTTCAGACCACATGACTTTTTCACAGAGTGGCCA 665

Qy 644 accacatcgtcacttgagggaacatttattatgccaccgttttttcaaaaaatgccttg 703
Db 666 ACTTCTCTTTTAATTTTGGAGCCCTATCTATTTTATTTTATTTTCAAGATGAAG 725

Qy 704 aaatagctctgaaattctccaaacacctgttacggagtgatgatctctacagcacat 763

```


QY 846 ctgcccacaggaagccgttgcctatggaattgaagcctacataatgaagctctctggaga 905
 Db 481 CTGTCATCAGGAAAGCCATTCCTATGGAATTTGAAGCCATACATTAATGCTTCGGAGA 540
 QY 906 acatggaattggtttctctctcttgggatacaatggtctcagaataatccagagaagaagc 965
 Db 541 ACATCGGAATTTGCTGTTCTCTTTGGGATCAATGCTCTCAGAAATTCAGAGAAAGAAAGC 600
 QY 966 latggaattgctgctgatttgggcaaaatccctcagacagctcctgtggcggtacactgg 1025
 Db 601 TATGCAATTTGCTGATGTTTGGGCAAAATCCCTCAGACAGTCTCTGFGCGGTACACTGG 660
 QY 1026 aaccgacacatgaaattcgttggaacaacaacagatacttgaagtgctaccaccaaaagc 1085
 Db 661 AACCCGACCATCAATCTTGGGCAACACAGATACCTTTAAGTGGCTACCCCAAAAGCA 720
 QY 1086 tctgttggtcacccgtagccctgctcttatcccccacatgctgttcccatggtgttta 1145
 Db 721 TCTGCTTGGTCACCCGATGACCCGCTGCTTTATCACCCATGCTGTTCCCATGGTGTGTTA 780
 QY 1146 tgaagcatatgcaatcgcttccatgctgagtgatgacccctgttggatgacagatgga 1205
 Db 781 TGAAGCATATGCAATGGGCTTCCATGGTGGATGATGCCCTGTGTTGGTGATCAGATGGA 840
 QY 1206 caatgcaagcgtatgagactaaggagcgtgagtgagcctggaatgattcttggaatgac 1265
 Db 841 CAATGCAAGCGCATGGAGCACTAAGGAGCTGGAGTGACCCCTGAATCTTCTGGAATGAC 900
 QY 1266 tctggaagatttagaagctctcaaaagcagtcacatgacaaagttacaagagaa 1325
 Db 901 TCTCTAGATTTAGAAATGCTCTAAAGGAGTCATCATGACAAAGATTACAGAGAA 960
 QY 1326 catatgcgctctccagccttcaagagcgcgcgtgagcgtgagcctggaactggccgt 1385
 Db 961 CATCATCGGCTCTCCAGCCTTCAAGAGCCGCCGCTGGAGCGCTGGACCTGGCCGT 1020
 QY 1386 gtctgggtgagttgtgtagagcacaagggcgcgcacacctgcccgcgcagccca 1445
 Db 1021 GTTCTGGTGGAGTTGTGATGAGCACAAGGGCGGCCACACCTGCGCCCGCAGCCCA 1080
 QY 1446 cgaactcaactggtacagatcaccattccttggaactgattggttctctcttggcgtgct 1505
 Db 1081 CGACCTCACCTGTGATCAGTACCATCTCTTGGACGTGATGGTGTCTCTTGGCCGCTGT 1140
 QY 1506 gtgacagtgcccttcatcacctttaaattgtgcttattgctaccggaaatgcttggtg 1565
 Db 1141 GCTGACAGTGGCTTCATCACCTTAAATGTTGCTTATGCTACCGGAAATGCTTGGG 1200
 QY 1566 gaaaagggcgagttgaagaaagcccaaaatccaaagaccattgagaagtggtggaa 1625
 Db 1201 GAAAAGGGCGAGTTAAGAAAGCCCAAAATCCAAGACCCATGAGAAGTGGGTGGAA 1260
 QY 1626 ataatgtaaaatttgaacattcctagctatttccaaacttgaacacagaatcagtg 1685
 Db 1261 ATAAGTAAAATTTGAACCATCTCTAGTCAATTTCCAAACTTGAACACAGATACAGTGT 1320
 QY 1686 taaattcatatttattttaaagaaatacttgcataaataatcagccccaagtgac 1745
 Db 1321 TAAATTCATTTTATTTTATTAGGAATACTTTTGCTTAATTAATCAGCCCGCAGATGC 1380
 QY 1746 tttaaaaaattctcttaataaaaaataatagactcgtagtcagtaagatatattgaata 1805
 Db 1381 TTTAAAAAATTTCTTTAAAAATAAATAATAGACTCGTAGTCAGTAAAGATATTGTAATA 1440
 QY 1806 tgaatcgccccctcgtgcttctgacagatgacatggtccatttttcagagagcg 1865
 Db 1441 TGTATCGTGGCCCTCTGCTGTCTTTGATCAGGATCAGATGTCCTATTTTCAGAGAGC 1500
 QY 1866 tgcagagagcgtggcattctagattactttctactctgaacacatgcccgtgtttggag 1925
 Db 1501 TGCAACAGAGGTGGCATCTAGATTACTTTTCTTACTTGAACATGATGCTGTTGGAG 1560
 QY 1926 tgcgggattcaaggttggtccccaccgctgcccactactgcaaatggcgttttaatttat 1985

Db 1561 TCGGGATTCAAAGGTGGTCCACGGCTGCCCTACTGCAAAATGGCAGTTTAAATCTTAT 1620
 QY 1986 ctttggctctcagagtggttgaattgaatccttaaccataatggtcagtcctcatct 2045
 Db 1621 CTTTGGCTCTCGAGATGGTGGCAATTTGATCTTAACCAATATGTTGTCCTCATCT 1680
 QY 2046 ctgctcgtctcattggtgcccacctgtgtgtttaagaagggaagcctttgtaccttag 2105
 Db 1681 CTGCTGCTTCATAGGTCACCTTGTGTGTGTTAAAGAGGAAGCTTTGTACCTTTAG 1740
 QY 2106 agtagtgtaaatgaatgactgcttgagtgacactgagacagcatatgattcttgc 2165
 Db 1741 AGTGTAGCTTAAATGANTGATGGCTTGGAGTGCACCTGAGAACAGCATATGATTTCTTGC 1800
 QY 2166 ttggggaaagaatgatgctatgaaattggtggtgtgtattttgagaagataatcat 2225
 Db 1801 TTTGGGAAAGAGATGATGCTATGAAATTTGGTGGTGTGTATTTCGAGAAGATAATCAT 1860
 QY 2226 tgcttatgtcaatgagctgaatttgataaaaccccaaaacacagctatgaagtgctgg 2285
 Db 1861 TGCTTATGTCAAATGGAGCTGAATTTGATAAAACCCAAATACAGCTATGAGTGCTGG 1920
 QY 2286 gcaagttactttttcttgatgttctctacaact 2320
 Db 1921 GCAAGTTTATCTTTTCTGATGTTTCTCTACAAC 1955

RESULT 10
 HUMHUGBR2 HUMHUGBR2 2368 bp mRNA linear PRI 07-MAR-1995
 LOCUS Human bilirubin UDP-glucuronosyltransferase isozyne 2 mRNA,
 DEFINITION complete cds.
 ACCESSION M57951.1 GI:184474
 VERSION UDP-glucuronosyltransferase; bilirubin UDP-glucuronosyltransferase.
 KEYWORDS Human adult female liver, cDNA to mRNA, clones Z17 and Z17P5E2.
 SOURCE Human adult female liver, cDNA to mRNA, clones Z17 and Z17P5E2.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 2368)
 AUTHORS Ritter, J.K., Crawford, J.M. and Owens, I.S.
 TITLE Cloning of two human liver bilirubin UDP-glucuronosyltransferase
 cDNAs with expression in COS-1 cells
 JOURNAL J. Biol. Chem. 266 (2), 1043-1047 (1991)
 MEDLINE 91093210
 FEATURES Location/Qualifiers
 1..2368
 /organism="Homo sapiens"
 /isolate="AK"
 /db_xref="taxon:9606"
 /map="Unassigned"
 /clone="Z17"
 /sex="female"
 /tissue_type="liver"
 /dev_stage="adult"
 /tissue_lib="lambda-ZAP"
 1..2368
 /gene="UGT2B"
 30..1634
 /gene="UGT2B"
 /standard_name="bilirubin UDP-glucuronosyltransferase
 isozyne 2"
 /EC_number="2.4.1.17"
 /codon_start=1
 /product="UDP-glucuronosyltransferase 2 family polypeptide
 B"
 /protein_id="AA063196.1"
 /db_xref="GI:184475"
 /db_xref="GDB:G00-127-753"
 /translation="MARGLOVPLPRLATGLLLLSVQPAESGKVLVVPDGSFNLISM
 REALRELHARGHQAVLTPENVNMHKEEFKFTLAYAVPMTQKDFRVTLYGTGTFE
 TEHLKRYSRSMALMNVSLALHRCVVELLHNEALIRHLNATSFQVLTDPVNLGAV

gene
 CDS

Qy 248 a

9	QY	cagttctctgatggctgtcacagggttgaccagccacctctctctatgtgtgtgtcgtc	68
25	Db	CACCTCTGGCATGTGGTTCACAGGTTGGACCAAGCCCCCTTCCCTATGTGTGTGTGTGCT	84
69	QY	gcctgacctgtggtctttcccaaggcaaggaaagctactggtagtcaccatggtggagacca	128
85	Db	CTGTGACCTGTGCTTTTGGCCAGACAGGGAAGCTAGTGTAGTCCCATGGATGGAGGCCA	144
129	QY	ctggtttcacatcaggtcgtgtgtgagaaacctcattctcagggggcatgagggtggttgt	188
145	Db	CTGTTTCACATFAGGTCGGTGTGGAGAAACTCATTTCTCAGGGGGCATGAGGTGGTTGT	204
189	QY	agttca-tgcacagaagttgaagttgcaactgggaagatcactgaattgcacagttgaagctt	247
205	Db	AGTCAGTGCACAGAGTGAGTTGGCACTGGGAAGATCAATGAATTGCACAGTGAAGACTT	264
248	QY	attcaactctatatacactggaggtactctggaccggaggttcaagggtcttttgcaccatgctc	307

DEFINITION Homo sapiens, Similar to UDP glycosyltransferase 1 family, polypeptide A9, clone MGC:9353 IMAGE:3850654, mRNA, complete cds.

ACCESSION BC020971

VERSION BC020971.1 GI:18088086

KEYWORDS MGC.

SOURCE human.

ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. 1 (bases 1 to 2423)

AUTHORS Strausberg, R.

TITLE Direct Submission

JOURNAL Submitted (03-JAN-2002) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA

REMARK NIH-MGC Project URL: <http://mgc.nci.nih.gov>

COMMENT Contact: MGC help desk
Email: gcaps-r@mail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)
Center, Stanford University School of Medicine, Stanford, CA 94305
Web site: <http://www-shgc.stanford.edu>
Contact: (Dickson, Mark) mcd@paxil.stanford.edu
Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers, R. M.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Series: IRAP Plate: 20 Row: m Column: 8
This clone was selected for full length sequencing because it passed the following selection criteria: GenomeScan gene prediction, Similarity but not identity to protein.
Location/Qualifiers
1. 2423
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="MGC:9353 IMAGE:3850654"
/tissue_type="Colon, adenocarcinoma"
/clone_lib="NIH_MGC_65"
/lab_host="DH10B"
/note="vector: pCMV-SPORT6"
70. .1662
/codon_start=1
/product="Similar to UDP glycosyltransferase 1 family, polypeptide A9"
/protein_id="AAH20971.1"
/db_xref="GI:18088087"
/translation="MARAGWTSPVPLCVCLLTGCFAPAGKLLVPMDSGHFTMQSV
VEKILRGHEVVMVPMVSWQLRSNCTVKYTSYSLTDQNRPMFPAHAKWAQY
QSIFSLMSSSGFLDLFFSHCRSLFNDRKLVLYLKSSDFDAFLDPFDCTGLIYAKY
FSLPVVFNRYGIFCHLHEEGAQCAPLSPVNDLIGFSDAMTEKRWVNHVHLEHDL
FCOYLFFNLEIASRLIOTPVAYDLYSHTSIWLLRTDFLDYKPKVPMNMFIFGGIN
CHQKPLPFEFAYINASGHEGLVPSLGSVSEIPEKKAMADALGKIPOTLWRY
TGRPSNLANNILYKWLPLQNDLLGHMTRAFITHAGSHGVYESICNGYPMWPLFG
DOMNARKMETKGVTLNLEMTSEDLNALKAVINDKSTKENIMKLSLKHDPVE
FLDLAVFEWFVWRHKGAPHLRPAADHLTWQYHSLDVIGFLAVLVITVAFITFCCA
YGRKCLGKGRVKAHKSHTH"
BASE COUNT 643 a 514 c 559 g 707 t
ORIGIN

Query Match 94.6%; Score 2194.6; DB 9; Length 2423;
Best Local Similarity 97.0%; Pred. No. 0;
Matches 2247; Conservative 0; Mismatches 69; Indels 1; Gaps 1;

QY 5 gctcagctctctgatgcttcacaggttgacaccccccctctctctatgtgtgtc 64
|||||

Db 56 GCTCAGCTCTCTGATGCTCGCGCAGGTGGACACGCCCGCTCTCTTATGTGTGTGTC 115
|||||

QY 65 tgetctgacctgtggtttgccagagggcagggaactactgtagtgcacatgatgga 124
|||||

Db 116 FACTGCTGACCTGTGGCTTGGCGAGGCGGGAAGCTGCTGCTAGTCCCATGGATGGA 175
|||||

QY 125 gccactgggttcacocatgagtcggtggtgagagaaactatctcaggggggcagtgaggtgg 184
|||||

Db 176 GTCACTGGTTTCCCATGCACTGGTGGTGGAGAACTTATCTCAGGGGGCATGAGGTGG 235
|||||

QY 185 ttgtatgcatccagaggtgagttggaactggaagagatcactgaattgcacagtgaga 244
|||||

Db 236 TTGTAGTCATCCAGAGGTGAGTTGGCAACTGGNAGATCAGTGAATTCACAGTGAAGA 295
|||||

QY 245 ctattcaacttcataacccctggaggtctggaccgggagttcaaggtctttggccatg 304
|||||

Db 296 CTTACTCAACTCTGCTACACTCTGGAAGATCAGAACCGGAATTCATGCTTTTCCCATG 355
|||||

QY 305 ctcaatgaaagacaagtacgaagtatatctctatttaagtggttcatacaatgaca 364
|||||

Db 356 CTCAATGGAAGACACAGGCACAAAGTATATTTCTCTATTAATGAGTTCATCCAGTGGTT 415
|||||

QY 365 ttttgacttatttttccaattgcaggagttgttttaaaagacaaaaaattagtagaat 424
|||||

Db 416 TTCTTGACTTATTTTTCGCATTGTCAGAGCTTTGTTTAAATGACCGAAAATTAGTAGAAT 475
|||||

QY 425 acttaaaagagagttcttttgatgaggtttctcgatcctctttgataactgtggtctaa 484
|||||

Db 476 ACTTAAAGAGAGATCTTTTGTATGACAGTGTCTTGGAATCTTTTGTACCTGCGCTTAA 535
|||||

QY 485 ttgttgcacaaatattctccctccctcctcgtgttcttcgccagggggaatacttgcac 544
|||||

Db 536 TTGTTGCTAAATATTTCTCCCTCCCTCTGCTGCTTCCACAGGGGAATATTTTGGCAC 595
|||||

QY 545 atctgaaagagtgacagtcctctcctctctctctctctctctctctctctctctctctag 604
|||||

Db 596 ATCTTGAGAGAGGTGACAGTGCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTAG 655
|||||

QY 605 ggttctcagatgcacttctcaaggagagagtaogaaacacacatcatgactgtgag 664
|||||

Db 656 GGTTCCTCAGATGCCATGACTTTCAAGAGAGAGATATGGAAACACATCTGTCGCTGGAG 715
|||||

QY 665 aacatttattatgcacgcgttttttcaaaaatgcctagaataatagcctctgaaattctcc 724
|||||

Db 716 ACATTTATTTGGCAGTATCTTTTGTAGAAATGCCCTAGAAATAGCCCTCTGAATCTCC 775
|||||

QY 725 aaacacctgttacggagtagatctctacagccacacatacttttgggtgttggaacgg 784
|||||

Db 776 AAACCCCTGTACGCGCATATGATCTCTACAGTCACACATCAATTTGGTGTGGCAACGG 835
|||||

QY 785 acttgttttgactatcccaaacccgtgatgcccaacatgactcttcatgttggtatca 844
|||||

Db 836 ACTTTGTTTGGACTATCCCAACCCGTGATGCCCAACATGATCTTCATTTGGTGTATCA 895
|||||

QY 845 actgcaatcaggggaagccctgctctatggaattgagcctacataatgcttctctgag 904
|||||

Db 896 ACTGTATCAGGGAAGCCCATTCCTCTATGGAATTTGAAGCTTACATTAATGCTTCTGGAG 955
|||||

QY 905 aacatgaaatgtgtttctctctgtggatcaatggtctcagaaattccagagagaagaa 964
|||||

Db 956 AACATGGAATTTGGTTTCTCTTTGGGATCAATGGTCTCAGAAATTCAGAGAAGAAAG 1015
|||||

QY 965 ctatggcaatgtcgtatgcttttgggcaaaatccctcagacagtcctgtggtgggtacactg 1024
|||||

Db 1016 CTATGGCAATTTGCTGATGCTTTTGGGCAAAATCCCTCAGACAGTCTCTGTTGGGGTACACTG 1075
|||||

QY 1025 gaaccgcacatcgaatcttcggaacaacaagatacttgttaagtggctacccccaaacy 1084
|||||

Db 1076 GAACCCGACCATGAATCTTGCGAACACACAGATACTTGTGTAAGTGGTGTACCCCAACG 1135
|||||

QY 1085 atctgttggtcacccatgacccgtgcctttataacccatgctgtgtccatgtgttt 1144
|||||

Db 1136 ATCTGTCTGGTCCACCGATGACCCGTGCTTATACCCATGCTGTTGCCATGTGTGTT 1195
|||||

QY 1145 atgaagcatatgaaatggcgtttcccatggtgatgagccctgtttgttgatcagatgg 1204
|||||

Db 230 TTCTAGTCTATGCCAGAGGTGAGTTGGCAACTGGGAAAAATCACTGAATTCACAGTGGAAGA 289
 Qy 245 cttattcaactcatataccctggaggtatctggaccgggagttcaaggcttttgcccatg 304
 Db 290 CTTTACTCAACCTCATACACTCTGGAGGATCTGGACGGGAATTCATGGATTCGCCGATG 349
 Qy 305 ctcaatggaaagacacaagttacagaatatattctctattatttggttttcatacaatgaca 364
 Db 350 CTCATGGAAGACACAAGTACGAGGTTTGTCTCTATTTCTGTAGTTTCATCCAAATGGTT 409
 Qy 365 tttttgacttatttttttcoaatgtaggagttgtttttaaagacaaaaattagtagaat 424
 Db 410 TTTTAAACTTATTTTTCGCATTCGACGAGTTTGTGTTTAAATGACCGAAAAATTAGTAAT 469
 Qy 425 acttaaaagagagttcttttgatgagttttctcaatcttttgataaacttggtgttaa 484
 Db 470 ACTTAAGGAGAGTCTTTTGAAGCGGTTTCTTGTATCTTTTGTATCCCTTGTGGCTTAA 529
 Qy 485 ttgttgcaaatattttctccctccctccgtggtcttcgccggggaatactttgccact 544
 Db 530 TTGTTGCCAAATATTTCTCCCTCCCTCTGTGTCTTCGCCAGGGGAATAGCTTGCCACT 589
 Qy 545 attctgaagaggtgacagtgccctgctctctcttctctatgctcccgagaattctctag 604
 Db 590 ATCTTGAAGAGGTGCACAGTGCCTGCTCTCTTCTCTATGTCCCCAGAAATCTCTTAG 649
 Qy 605 ggttctcagatgcactgaactttcaaggagagagtagcggaaccacatcatgcaacttgagg 664
 Db 650 GGTCTCTAGATGCCATGACTTTCAAGGAGAGAGTAGTACGGAACACATCATGCACCTTGAGG 709
 Qy 665 aacattattatgcccacgcttttttcaaaaatgccctagaaaatagcctctgaaattctcc 724
 Db 710 AACATTTATTTTGGCAGTATTTTCCAAAAATGCCCTAGAAATAGCTCTGAATTCCTCC 769
 Qy 725 aaacacctgttagagtagatgctctacagcaacatcaaatgttggttgtagaacg 784
 Db 770 AAACACCTGTCAAGCATATGATCTCTACAGCCACACATCAATTTGTTGTTGCGAAGAC 829
 Qy 785 atttgttttggactatcccaaacccgtgatgcccaacatgatcttcaattggtggtatca 844
 Db 830 ACTTTGTTTGGACTATCCCAACCCGTTGATGCCCAATATGATCTTCATTGGTGTATCA 889
 Qy 845 actgcatacaggaaagccgttgctctatggaatttgaagcttaagctcaataatgcttgag 904
 Db 890 ACTGCCATCAGGAAAGCCATTCCTATGGAATTTGAAGCTTACATTAATGCTCTGGAG 949
 Qy 905 acataggaaattggttttctctcttgggatacaatggtctcagaataattccagagaagaag 964
 Db 950 AACATGGAAATGTGTTTCTCTTTGGGATCAATGGTCTCAGAAATTCAGAGAGAAAG 1009
 Qy 965 ctatggcaattgctgagcttgggcaaaatccctcagacagctctgtggcggtacatg 1024
 Db 1010 CTATGSCAATGTGATGCTTTGGGCAAAATCCCTCAGACAGCTCTGTGGGGGTACACATG 1069
 Qy 1025 gaacccgacatcgatctctgcgaacacacgatacttgaagtggtctaccccaaaag 1084
 Db 1070 GAACCCGACCATCGAATCTTGGCAACACACAGATACTTTAAAGTGGCTACCCCAAAAGC 1129
 Qy 1085 attctgttgggtcaccgtagcccggtcttcttaccacccatgctggttcccatggtttt 1144
 Db 1130 ATCTGCTGTGTCACCCGATGACCCGTGCTTTATACCCCATGCTGGTGTCCCATGGTGT 1189
 Qy 1145 atgaagcatatgcaatggcgttcccatggtgatgagtccttgggttgatcagatgg 1204
 Db 1190 ATGAAGCATATGCAATGGCGTTCOCATGGTATGATGATGCCCTTGTGTTGGTATCAGATG 1249
 Qy 1205 acaatgcaagcgcatggagacataaggagctgagtagtaccctaaatggttctggaataa 1264
 Db 1250 ACAATGCAAGCGCATGGAGACTAAGGAGCTGGAGTGGACCTGAATGTTCTGGAATGA 1309
 Qy 1265 ctctcgaagatttagaaaaatgctctaaagcagtcacatgacaaaagttaacagggaga 1324
 Db 1310 CTTTCTGAAGATTTAGAAAAATGCTCTAAAGCAGTCAATCAATGACAAAAAGTTACAAGAGA 1369

Qy 1325 acatcatgcgctctccagccttcaagaccccggtgagccgctgagccgctgagccgctgagccg 1384
 Db 1370 ACATCATGCGCTCTCGCGCTTCACAGGACCCCGGTGGAGCGCTGGACCTGGCCG 1429
 Qy 1385 tgtcttggtgaggtttgtgtaggacaaaggcgccacacacctgcgcccccgagccc 1444
 Db 1430 TGTCTGGGTGGAGTTTGTATGAGGCACAAGGGCGCGCACACCTGCGCCCGCAGCCC 1489
 Qy 1445 aqacactcaccctggtacagtagtaccattccttgagcgtgattggttctctctggtcgctg 1504
 Db 1490 ACGACCTCACCTGTCACAGTACCATTCCTTGGACGTGATTGGTTTCTCTTGGCGGTGG 1549
 Qy 1505 tgcctgacagtgcctctcaatcaacttaaatgtgtgcttatggtctaccgcggaatgcttg 1564
 Db 1550 TGTCTGACAGTGGCGCTCATACACTTTTAAATGTTGTGCTTATGGCTACCGGAAATGCTTGG 1609
 Qy 1565 ggaaaaaagggcgaggttaagaaagcccaaaatccaagaccattggaagtggtgggga 1624
 Db 1610 GGAATAAAGGGCGAGTTAAGAAAGCCCAAAATCCAAAGACCCATTTGAGAGTGGGTGGGA 1669
 Qy 1625 aataaggttaaaatttgaaccattcccttagtcaatttccaaacttgaaaaacagaatcagtg 1684
 Db 1670 AATAAGGTAAAAATTTGAACCAATTCCTTAGTCAATTTCCAAACTTGAAAAACAGAAATCAGTG 1729
 Qy 1685 ttaattccttttattcttatttaaggaataactttgcatataaatttaactcagccccagagtg 1744
 Db 1730 TTAATTTCAATTTATTTTATTATTAAGGAAATTAATTTGCATAAATTAATCAGCCCCAGAGTG 1789
 Qy 1745 ctttaaaatctctcttaataaaaaataatagactcgctagtcagtaagaaagatttgaat 1804
 Db 1790 CTTTAAAAAATCTCTTAATTAATAAATAATAGACTGCTAGTACGATAAAGATATTGAAAT 1849
 Qy 1805 atgtatcglgccccctccggtgcttttgatcagagtgacatgctccatttttcagaggac 1864
 Db 1850 ATGTATCTGTGCCCGCTCCCGTCTTTTGTATCAGGATGACATGTGCATTTTTCAGAGGAC 1909
 Qy 1865 gtgcagacggcggtgcatcttagattactttctctctctgaaacatggtcgttggga 1924
 Db 1910 GTGCAGACAGCGTGGCATCTTAGATTACTTTCTTCTCTGAAACATGGCTGTTTGGGA 1969
 Qy 1925 gtgcggatgtaaaaggtggtccccccgctgccccactgcaaatggcagttttaaactta 1984
 Db 1970 GTGCGGGATTCAAAGGTGTGCCACCGCTGCCCTACTGCAATGGCAGTTTAAATCTTA 2029
 Qy 1985 tcttttgctctctgcagatggttgcaattgac-cttaaccaataatggtcagctcctat 2043
 Db 2030 TCTTTTGGCTTCTGCAGATGTTGCAATTTGATCTCTTAACCAATTAATGGTCAGTCTCAT 2089
 Qy 2044 ctctgctccttcaatagtgccacctgtgtgtttaaagaaggagagctttgtaccttt 2103
 Db 2090 CTCTGCTCTGCTTCTATAGGTGCCACTTGTGTGTTTAAAGAGGGAAGCTTTGTACCTTT 2149
 Qy 2104 agagtgtagtgaaatgaatgaatggtgtgagtgactgagaaacagcatatgattctt 2163
 Db 2150 AGAGTGTAGTGAATGAATGAATGGCTTGGAGTGCACGTGAGAACAGCATATGATTTTT 2209
 Qy 2164 gctttgggaaaaagaatgactatgaaattggtgggtggttatittgagaagaataatc 2223
 Db 2210 GCTTTGGGAAAAAGAAATGATGCTATGAAATTTGGTGGGTGTATTGAGAAAGATAATC 2269
 Qy 2224 attgcttatgtcaaatggagctgaatttgataaaaccccaaaacacagctatgagtgct 2283
 Db 2270 ATTGCTATGTCAATGGAGCTGAATTTGATAAAACCCCAAAATACAGCTATCAAGTGCT 2329
 Qy 2284 gggcaaatgttcttttttctgatttctcacaact 2320
 Db 2330 GGGCAAGTACTTTTTTTTTTGTATGTTTCTCTCAACT 2366

RESULT 6
 BC020971
 LOCUS

BC020971 2423 bp mRNA linear PRI 22-JAN-2002

QY 1205 acaatgcaagagcatgagactaaaggagctgagtgagcctgaatgtctctggaatga 1264
|||||
Db 1220 ACAATGCAAAAGCGCATGAGACTAAGGAGCTGGAGTGACCCCTGAATGTTCTGGAATGA 1279

QY 1265 ctctcgaagattagaataatgctcttaaaagcagtcataatgacaaaagttaacagaga 1324
|||||
Db 1280 CTCTGGAATTTAGAAATGCTCTAAAGCAGTCATCAATGACAAAATTTACAAGAGA 1339

QY 1325 acatcagtcgctctccagccttcaaaagaccgcccgggtggagcgcgtggacacctggccg 1384
|||||
Db 1340 ACATCATGCGCTCTCCAGCTTCAACAGACCGCCGGTGGAGCGCTGGACCTGGCCG 1399

QY 1385 tggcttggtgagttgtgatgagcaaaagggcgccacacacctggcccccacccc 1444
|||||
Db 1400 TGTTCTGGGTGGAGTTGTGATGAGGCAACAGGGGGCGCCACACCTGCGCCCGGACGCC 1459

QY 1445 acgacctcacctgggtacagctaccattccttgagcgtgattggttctctctggccgtcg 1504
|||||
Db 1460 ACGACCTCACCTGGTACCAGTACCATTCTTGGACGCTGATGGTTCCTCTTGGCCGTCG 1519

QY 1505 tgcgcagctgacctcacccttaaatgtgtgtcttatgctaccggaatgcttg 1564
|||||
Db 1520 TGCTGACAGTGCCCTTCATCACCCTTAAATGTTGTGCTATGCTACCGGAATGCTTGG 1579

QY 1565 ggaataaaagggcgagtttaagaaagcccaaaatccaaagaccattgagaagtggtggga 1624
|||||
Db 1580 GGAATAAAGGGCGAGTTAAGAAAGCCACAAAATCCAGACCCATTGAGAAGTGGGTGGGA 1639

QY 1625 aataaagtaaaatttgacattccctcagctcatttccaaactgaaacagaaatcagtg 1684
|||||
Db 1640 AATAGGTAAATTTTGAACCATTCCTTGTCTTCTTCAACCTTGAAACAGATCATGTG 1699

QY 1685 ttaattcatcttattcttatttaaggaatacttgcataataatcagccccagagtg 1744
|||||
Db 1700 TTAATATTCATTTTATTTTATTAAGGAATACTTTGCATAAATAATCAGCCCCAGAGTG 1759

QY 1745 ctttaaaataatcttaataaaataagactcagctcagtaagaaatattgaat 1804
|||||
Db 1760 CTTTAAATAATTCCTTAATAATAATAATAGACCTGCTGATGATGATTAAGATATTTGAAT 1819

QY 1805 atgtatgctgccccctccggtgtcttttgatcaggatgacatgtgccatttttcagagagac 1864
|||||
Db 1820 ATGTATGCTGCCCCCTCCGGTGTCTTTGTATCAGGATGACATGTGCCATTTTTCAGAGGAC 1879

QY 1865 gtgcagcagctggcattctagctacttcttacttacttgaacacatgacctgtttggga 1924
|||||
Db 1880 GTGCAGACAGCTGGCATTCTAGATTACTTTTCTTACTCTGAAACATGCCCCGTTGGGA 1939

QY 1925 gtgcggattcaaaagtggtgcccaccgctgccccctactgcaaatggcagttttaaattta 1984
|||||
Db 1940 GTGCGGATTTCAAAGGTGGTGGTCCACCGCTGCCCTTACTGCAAAATGGCAGTTTTTAATCTTA 1999

QY 1985 tctttgctctcagatggttgcaatgatccttaacaaataatggcagctcctcctc 2044
|||||
Db 2000 TCTTTTGGCTCTGAGATGTTGCAATTTGATCTTACCANTANTGGTCAGCTCTCATC 2059

QY 2045 tctctcctctcagtagtgcacacctgtgtgttttaagaaggaagccttbtgacacctta 2104
|||||
Db 2060 TCTGCTCTGCTTCATAGGTGCCACCTGTGTGTTTTAAAGAGGAGGCTTTGTACCTTTA 2119

QY 2105 gagttagtgaaatgaatgaatgagctgagtgacagcagacacacatgatttcttg 2164
|||||
Db 2120 GAGTGTAGTGAATGAATGAATGGCTTGGAGTGCATGAGACACATATGATTTTTTG 2179

QY 2165 ctttgggaaaaaatgatgctatgaattggtgggtgggtgtatttgagaagataatca 2224
|||||
Db 2180 CTTTGGGAAAAAAGATGATGCTATGAAATTTGGTGGTGTATTTTGAAGAAGATAATCA 2239

QY 2225 ttgcttatgtcaaatggagctgaatttgataaaaccccaaaatcacagctatgaatgctg 2284
|||||
Db 2240 TTGCTTATGTCAAAATGGAGCTGATTTGTATAAACCCCAAAATACAGCTATGAGTGTG 2299

QY 2285 ggcaagtttaactttttttctgtatgttttctactaaact 2320
|||||
Db 2300 GGCAAGTTTACTTTTTTTTGTGATGTTCTCTCAACT 2335

RESULT 5
AF462268 2394 bp mRNA linear PRI 07-FEB-2002
LOCUS Homo sapiens UDP-glucuronosyltransferase UGT1A8*1 (UGT1A8) mRNA,
DEFINITION UGT1A8*1 allele, complete cds.
ACCESSION AF462268
VERSION AF462268.1 GI:18568218
KEYWORDS human.
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 2394)
AUTHORS Huang,Y.-H., Nguyen,N., Geske,D. and Tukey,R.
TITLE Identification and functional characterization of UGT1A8*1,
UGT1A8*2 and UGT1A8*3
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 2394)
AUTHORS Huang,Y.-H., Nguyen,N., Geske,D. and Tukey,R.
TITLE Direct Submission
JOURNAL Submitted (21-DEC-2001) Pharmacology, UCSD, 9500 Gilman Drive, La
Jolla, CA 92093, USA
FEATURES
Location/Qualifiers
1..2394
/organism="Homo sapiens"
/db_xref="taxon:9606"
/tissue_type="colon"
1..2394
/gene="UGT1A8"
/allele="UGT1A8*1"
64..1656
/gene="UGT1A8"
/codon_start=1
/product="UDP-glucuronosyltransferase UGT1A8*1"
/protein_id="A175964.1"
/db_xref="GI:18568219"
/translation="MARTGWTSPICVSLITCGFAEAGKLLVPMDSHFWQSV
VERLIRGHEVVVMPVEVSWQLGKSLNCTVYTSYTLDELDFEMDFADQWRAQV
RSFLFLSSNGFNFLFHSRSLFNDRKLVEYLFKESFADFLDFDACCGLIYAKY
FSLPVSFVARGIACHYLEEQAQCPAPLYSPRIILGFSDFAMTKFVRNHHMLBEHL
PCOYFSKNALEIASIQLPTVATLYSHTLSWLLRDFVLDYPKPVMNMFIFGIN
CHQCKPLPMEFEAYINASGEHGIWFSLSGWSVEIPEKKAMADALGKIPQIVLMRY
TGTPSNLANNTILVKLPQNDLIGHPTAFITHAGSHGVYESICNGVPMVMPFLFC
DQMDNAKMETKAGVYTLNLEMTSEDLNALKAVINDKSKYKNIHLRLSLHKKRPE
FLDLAVFWFEFVRHKGAPHLRPAADHLTWYQYHSLDVIGFLLAVLVITFAITFKCCA
YGYRKCGLGKRKVRKKAHKSHTH"
BASE COUNT 624 a 513 c 555 g 702 t
ORIGIN

Query Match 95.1%; Score 2207.4; DB 9; Length 2394;
Best Local Similarity 97.3%; Pred. No. 0;
Matches 2255; Conservative 0; Mismatches 61; Indels 1; Gaps 1;

QY 5 gctgcagttctctgagcttgacaggggtggacccccccttctctatgtgtgtc 64
|||||
Db 50 GCTGCAGTTCTCTCATGCTCGCACAGGTGGACCCCATTCCTTATGTTTCTC 109

QY 65 tgcgtgcacctgtgctgttcgcagaggaaggaagcactgtagtgcacgatgga 124
|||||
Db 110 TGCTGCTGACCTGTGGCTTGTCTGAGGAGGAGGAGCTGCTGGTAGTGCCTATGATGGGA 169

QY 125 gccactggttcaccatgaggtcggtgtgtggaagaaactcattctcaggggcatgagtg 184
|||||
Db 170 GTCACGTGTTACCATGACGTGCGTGGTGGAGAAACTTATCCTCAGGGGCGCATGAGGTG 229

QY 185 ttgtagtcatgcacaggtgagttgcaactggaagatcacgtgaattgcacagtgaga 244
|||||

QY	2161	cttgcttggggaaagaatgatgctatgaattgggtgggtgattttgagaagata	2220	Db	80	TGCTGCTGACCTGTGGCTTGGTGTGAGCAGGGAAGCTGCTGTAGTGCCTCATGGATGGGA	139	
Db	2161	CTTGCTTTGGGAAAAAGATGATGCTATGAATTTGGTGGTGGTGTATTTGAGAAGATA	2220	QY	125	gccactggttcacccaagaagtcggtggtggagaactcattctcagggggcattgaggtgg	184	
QY	2221	atcattgcttattgcaaaatggagctgaatttgataaaaccccaaaatacacagctataaot	2280	Db	140	GTCATGCTGCTCACCCTGCTAGTGGTGGTGGAGAACTTATCTCTCAGGGGGCATGAGGTGG	199	
Db	2221	ATCATGCTTATGTCAAATGGAGCTGAATTTGATTAACCCCAAAATACAGCTATGAAGT	2280	QY	185	tgtgtagtcattcagcagaggtgagttggcaactgggaagatcactgaattgacacagtgaaaga	244	
QY	2281	gctgggcaagttaacttttttctgattgtttccctacaact	2320	Db	200	TTGTAGTCAATGCCAGAGGTGAGTTGGCAACTGGGAAAAATCACTGAATTTGCACAGTGAAGA	259	
Db	2281	GCTGGCAAGTTTACTTTTCTGATGTTTCTCTACAACT	2320	QY	245	cttattcaacttcatactcaggaatctggaacccggaggttcaaggcttttgcacatg	304	
RESULT	4			Db	260	CTTACTCAACCTCATACATCTCGAGGATCTGGACCGGAATTCATGATTTGCCGATG	319	
LOCUS	AF462267	2363 bp	mRNA	linear	QY	305	ctcaatgggaagcaccaagcacgaagtatatattctcttataatgggtttccatacaatgaca	364
DEFINITION	Homo sapiens UDP-glucuronosyltransferase UGT1A8*2 (UGT1A8) mRNA,			Db	320	CTCAATGGGAAGACACAAGTACGAAGTTTGTCTCTATTTCTGAGTTTCATCCCAATGGTT	379	
ACCESSION	AF462267			QY	365	tttttgacttatttttcaaatgacagagttgttttaagacacaaaaattagtagaat	424	
VERSION	AF462267.1	GI:18568216		Db	380	TTTTTAACTTAATTTTTCGCATTCAGGAGTTGTTTAATGACCGAAAAATAGTAGAAT	439	
KEYWORDS	human.			QY	425	acttaagcagaggtcttttgatgagtggtttcttcgatcccttttgataaacttgggttaa	484	
ORGANISM	Homo sapiens			Db	440	ACTTAAAGAGAGGTTCTTTTGTATGCGGTGTTCTTGTATCTTTTGTATGCTTTCATGCTTAA	499	
REFERENCE	1 (bases 1 to 2363)			QY	485	ttgttgccaaatatttctccctccctccctcgtggtcttccagggggaatactttgacct	544	
AUTHORS	Huang, Y.-H., Geske, D., Nguyen, N. and Tukey, R.			Db	500	TTGTTGCCAAATATTTCTCCCTCCCTCTGTCGTCCTGCCAGGGGAATAGTGGCACT	559	
TITLE	Identification and functional characterization of UGT1A8*1, UGT1A8*2 and UGT1A8*3			QY	545	atcttgaagaaggtgcacagtgccctcctcctctctctctctctctctctctctctctctctct	604	
JOURNAL	Unpublished			Db	560	ATCTTGAAGAAGTGCACAGTGCCCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	619	
REFERENCE	2 (bases 1 to 2363)			QY	605	ggtctcagatgcacttcaagagagagtagtgcgaacacacatcatgcacttggagg	664	
AUTHORS	Huang, Y.-H., Geske, D., Nguyen, N. and Tukey, R.			Db	620	GGTCTCAGATGCCATGACTTTCAAGGAGAGAGTACGGAACCCACATCATGCACCTTGAGG	679	
TITLE	Direct Submission			QY	665	aacatttatatgccacgcttttttcaaaaatgccctagaaatagcctctgaaattctctcc	724	
JOURNAL	Submitted (21-DEC-2001) Pharmacology, UCSD, 9500 Gilman Drive, La Jolla, CA 92093, USA			Db	680	AACATTTATTTTGGCAGTATTTTTCAAAATGCCCTAGAAATAGCCTCTGAAATTCCTCC	739	
FEATURES	Location/Qualifiers			QY	725	aaacacctgttacggagtagtctctacagccacacataaatttggttgggttcgcaacgg	784	
source	1..2363			Db	740	AAACACCTGTACAGCATATGATCTCTACAGCCACACATCAATTTGTTGTTGTCGGAACAG	799	
gene	/organism="Homo sapiens"			QY	785	actttgtttggactatcccaaaccttgatgcccaacatgatctctctctctctctctctctctct	844	
	/db_xref="taxon:9606"			Db	800	ACTTTGTTTGGACTATCCCAACCCGTTGATGCCCAATATGATCTTCTATTGGTGGTATCA	859	
	/tissue_type="colon"			QY	845	actgccatcaggggaaagccgttgccctatggaatttgaagccctacataatgctctctgag	904	
	1..2363			Db	860	ACTGCCATCAGGGAAGCCATTTGCTATGGAATTTGAAGCCCTACATTTAATGCTTCCTGGAG	919	
	/gene="UGT1A8"			QY	905	aacatggaattggttct	964	
	/allele="UGT1A8*2"			Db	920	AACATGGAATTTGTGGTTTCTCTTTTGGGATCAATGCTCTCAGAAATTTCCAGAGAAGAAG	979	
	34..1626			QY	965	ctatggcaattgctgattgctttgggcaaaaatccctcagacagctcctgtggcggtacactg	1024	
	/codon_start=1			Db	980	CTATGGCAATTTGCTATGCTTTTGGCAAAAATCCCTCAGACAGTCTCTGTGGCGGTACACTG	1039	
	/product="UDP-glucuronosyltransferase UGT1A8*2"			QY	1025	gaacccgaccatcgaatcttgcgaacacacacatacttgaagtggtcaccaccccaaacg	1084	
	/protein_id="AAL75963.1"			Db	1040	GAACCCGACCATCGAATCTTGGGAACCAACAGATACTTGTAAAGTGGCTACCCCAAAACG	1099	
	/db_xref="GI:18568217"			QY	1085	atctgcttggttcaaccccgatgacccgttccctttatcacccatgctgggttcccatggtgttt	1144	
	/translation="MARTGWTSPILCVSLLLTCGFAEAGKLLVVPMDGSHWFTMQSV			Db	1100	ATCTGCTTGTCTACCCGATGACCGTCCCTTTATACCATGCTGTTCCCATGGTGT	1159	
	VEKLIRGHEVVVMPVSNOLGSLNCTVKTSTSYLTEDLDFEPMDFADQWKAQV			QY	1145	atgaagaacatgaatgcttcccatgataatcccttctgttggatgacagtg	1204	
	RSLSIFLSSNGFNLFPSHCRSLFNDRKLVEFKESFDFVLPDPACGLIVAKY			Db	1160	ATGAACCAATGCAATGGCTTCCCATGTTGATGCCCTTGTGGTATGATGATGG	1219	
	FLPSVWFRAGTCHLEGAQCPAPUSVPRILLGFSAMTFKERVNHIMLEHL							
	FCQYFSKALIEAIEILQPTVAYDLYSHTSIMLLRTDFVLDPKPVPNNHIFIGIN							
	CHQKPLMEFEAYINAGEHGIVVFSIGMSVEIPEKKAMAIAADLKIPQTLWRY							
	TTREPSLNANTILVKLPQNDLGHMPTRAFITHAGSHGVYESICNGVPMWVPLFG							
	DQMDNAKMETKAGVTLNVLNEMTSEDLENALKAVINDKSKENIMRLSLHLDREVE							
	PGLDAVFWFVVRHKGAPHLRPAHDLITWYQVHSLDVLGFLAVLTVAFITFKCA							
	YGYRCLGKGRVKKAKSKTH"							
BASE COUNT	617 a 503 c 551 g 692 t							
ORIGIN								
Query Match	95.6%	Score 2218.4	DB 9	Length 2363				
Best Local Similarity	97.4%	Pred. No. 0						
Matches 2255:	Conservative	0	Mismatches	61	Indels	0	Gaps	
QY	5	gctgcagttctctgagcttgcacaggggtgacacagccctctctctctctctctctctctctctct	64					
Db	20	GCTGCAGTCTCTCATGCTCGCACAGGGTGGACCAAGCCCATTTCCCTATGTTGTTCTC	79					
QY	65	tgtctgctgacctgtggttcccgagcaggggaagcactactgtagtgcacatggtggga	124					

Matches 2320; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
Qy	1	ctcagctgagttctctgagtgcttgacaggggtggaccagcccccctctcctctatgtgtg	60
Db	1	CTCAGCTGCAGTCTCTGATGGCTTGACAGGGTGGACAGGCCCTCTCTATGTGTG	60
Qy	61	tgtctctgctgaacctgtgcttgcgagggcaggggaagctaactagtgagtcacatgagat	120
Db	61	TGTCCTGCTGCTGACCTGTGGCTTTGCCAGGCGAGGAAGCTACTGTGTAGTCCCATGGAT	120
Qy	121	ggagagcaactggttcacacatgagtgctgggtggagaaactcattctcagggggcatgag	180
Db	121	GGGAGCCACTGGTTCCACCATGAGTGGGTGGTGGAGAACTCATTTCTCAGGGGCGATGAG	180
Qy	181	gtggttgtagtccagaggttgagtggcaactgggaagatcactgaatgcaagt	240
Db	181	GTGGTGTGTAGTCATGCCAGAGGTGAGTTGGCAACTGGGAAGATCATGAATTCACAGTG	240
Qy	241	aagactattcaacttcataccctggagatctggaccgggagttcaagggtctttggc	300
Db	241	AAGACTTATTCAACTTCATATACCTGGAGGATCTGGACCGGGAGTTCAAGGCTTTTGCC	300
Qy	301	catgctcaatggaaagcaagtaacgaagtatatctctatattgggttcatacaat	360
Db	301	CATGCTCAATGGAAAGCACAGTACGAAGTATATTTCTCTATTAATGGGTTCATACAAT	360
Qy	361	gacattttgacttatttttcaaatgacagagttgttttaagacaaaaaattagta	420
Db	361	GACATTTTGGACTTATTTTTCAAATTCAGGAGTTGTTTAAAGACAAAAAATTAGTA	420
Qy	421	gaatacttaaaagaggttcttttgatgcagtggtttctcgatctcttttgataaactggc	480
Db	421	GAATACTTAAAGAGAGGTCTTTTGATGCAGTGTCTTCGATCCTTTTGTAATCTGTGGC	480
Qy	481	ttaattgtgcaaatattctccctccctccgtggtcttcgccaggggaaactcttgc	540
Db	481	TTAATTTGTGCCAAATATTTCTCCCTCCCTCCGTGGTCTTCGCCAGGGGAATACTTTGC	540
Qy	541	caatacttgaagaagtgacagtgccctcctcctctctctctctctctctctcctcctc	600
Db	541	CACATCTTGAAGAAGGTGCACAGTCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	600
Qy	601	ttaggttctcagatgcctacttcaaggagagtagcggaaccacatcatgacttg	660
Db	601	TTAGGTTCTCAGATGCCACTTTTCAAGGAGAGAGTAGCGGAACCATCATGCACATTG	660
Qy	661	gagaaactttatgacacgttttttcaaaaatgccatgaagaatgcctctgaaatt	720
Db	661	GAGGAACATTTTATATGCCACCGTTTTTTCAAAAATGCCCTAGAAATAGCCCTCTGAAT	720
Qy	721	ctcaaacacctgttaoaggagtagatctctacagccacacatcaatttgggtgttcgga	780
Db	721	CTCCAAACACCTGTACGGAGTATGATCTCTACAGCCACACATCAATTTGGTGTTCGGA	780
Qy	781	acgaactttgttgactatccaaacccgtgagtcgcaacatgactctcatttggttgg	840
Db	781	ACGACTTGTGTGTGGACTATCCAAACCCGATGCCCAACATGATCTTCATGTGGTGT	840
Qy	841	atcaactgcacatcagggaagccgttgctatggaaatttgaagccctacaataatgctct	900
Db	841	ATCAACTGCCATCAGGGAAGACCGGTTGCCATGTGAATTTGAAGCCCTACATTAATGCTCT	900
Qy	901	ggagaacatggaattgtgtttctctttgggaatcaatgtctcagaatccagagaag	960
Db	901	GGAGAACATGGGAATTTGTGGTTTTCTTTTGGGATCATGTGTCAGAAAATTCAGAGAAG	960
Qy	961	aaagctatggcaattgctgagtcttgggcaaaaatccctcagacagctcctgtggcggtac	1020
Db	961	AAAGCTATGGCAATTTGCTGATGCTTTTGGGCAAAATCCCTCAGACAGTCTGTGGCGGTAC	1020
Qy	1021	actggaacccgacatcaatcttgcgaacacacagatactgttaagtggcctaccccaa	1080
Db	1021	ACTGGAACCCGACCATCGAATCTTGTGGAACACACAGATACCTTTGTAAGTGGCTACCCAA	1080
Qy	1081	aacgatctgttgtaaccctgacccgtgacccgtgctcttatcaaccatgctgggtcccatggt	1140
Db	1081	AACGATCTGCTTGGTACCCGATGACCGGTGCCCTTATCACCCATCTGTTGCCATGTT	1140
Qy	1141	gtttatgaaacatcatgcaatgcttcccatgctgatgcccttgttgggtgacag	1200
Db	1141	GTTTATGAAACATATGCAATGGCTTCCATGGTGATGATGCCCTTGTTTGGTGATCAG	1200
Qy	1201	atggacaatgcacagcgcatgagagtaaggagctgagtgagtgacctgaatgttctggaa	1260
Db	1201	ATGGACAATGCAAAAGCGCATGGAGACTAAGGGAGCTGGAGTACCCTGAATGTCTTGAA	1260
Qy	1261	atgactctgaagatttaaaaaatgctctaaagcagtcacatcaatgacaaagtacaag	1320
Db	1261	ATGACTCTGAAGATTTAGAAAATGCTCTAAAGCAGTCATCAATGACAAAAGTTACAG	1320
Qy	1321	gagacatcatgcctctccagccttcacagagcgcgcgcgtggagccgtggaacctg	1380
Db	1321	GAGACATCATGCGCTCTCCAGCCTTCACAAGGACGCCGCCGTGGAGCGCTGGACCTG	1380
Qy	1381	gcctgttctgggtgaggtttgtgatgagcacaaaggcgccacacctgcgcccgca	1440
Db	1381	GCCGTCTCTGGGTGGAGTTTGTGATGAGGCACAAGGGCGCGCCACACCTGCGCCCGCA	1440
Qy	1441	gccacgacctcaactggtaccagtcaccattccttggagcgtggttcttctctgtgccc	1500
Db	1441	GCCACGACCTCAGCTGGTACCAGTACCATTCTTGGACGTGATTGGTTCTCTCTTGCC	1500
Qy	1501	gtcgtctgacagtgcccttcacacctttaaagtgtgtcttatggtaccgcggaatgc	1560
Db	1501	GTCGCTGCACAGTGGCTTTCATCACCCTTAAATGTTGTCTTATGGCTACCGGAAATGC	1560
Qy	1561	tggggaaaaagggcgagtgaaagaagcccaaaatccaagaacccattgagaagtgggt	1620
Db	1561	TTGGGAAAAAGGGCGAGTTAAGAAGCCCAAAATCCAAGACCCATTGAGAAGTGGGT	1620
Qy	1621	ggaaataaggttaaaatttgaaccattccctctagtcatttccaaacttgaacacagatc	1680
Db	1621	GGAAATAAGGTAAATTTTGAACCATTCCTCTAGTCATTTTCCAACTTTGAAACAGAAATC	1680
Qy	1681	agtgtaaatcattcttcttatttaaggaaaaactttgcataataataatcaagcccgag	1740
Db	1681	AGTGTAAATTCATTTATTTCTTTTAAAGAAATACCTTTGCATAAATTAATCAGCCCGAG	1740
Qy	1741	agtgctttaaaaaatctctttaaataaaaaataatagactcgctagtcagtaaaagattt	1800
Db	1741	AGTGTCTTAAAAAATCTCTTAAATAAAAAATAATAGACTCGCTAGTCAAGATAATTT	1800
Qy	1801	gaatatgtatcggtcccccctccggtgtctttgatacaggatgacatgtgccattttcaga	1860
Db	1801	GAATATGTATCGTGCCCCCTCCGGTGTCTTTGATCAGGATGACATGTGCCATTTTTCAGA	1860
Qy	1861	ggagtgacagacagctggcattctctagattcttcttactctgaacatggcctgttt	1920
Db	1861	GGAGTGCAGACAGCTGGCATTTCTAGATTACTTTCTTACTCTGAACATGGCTGT	1920
Qy	1921	ggagtgccggatccaagtggtccaccgcgtccctactgcaaatggcagttttaat	1980
Db	1921	GGGAGTCCGGGATTCAAAGGTGTCCTCCACCGTCCCTACTGCAATGGCAGTTTAAAT	1980
Qy	1981	cttatcttttggctctgcagatgggttgcaattgactcttaaccataatggtcagtcct	2040
Db	1981	CTTATCTTTGGCTCTGCAGATGGTTGCAATGATCCTTAACCAATAATGTCAGTCT	2040
Qy	2041	catctctgctgtctcaatgagtgccacctgtgtgtttaaagaaggaagctttgtacc	2100
Db	2041	CATCTCTGCTCCGTCTCATTAAGTGGCACCTTGTGTGTGTTTAAAGAGGAAGCTTTGTACC	2100
Qy	2101	tttagagtgtagtgtaaatgaaatgaaatgcttggagtgactgagacacagcatatgattt	2160
Db	2101	TTTAGAGTCTAGTGAATGAATGAATGCTTGGAGTGCACCTGAGAACAGCATATGATT	2160


```

|||||
Db 1861 GGACGTGCAGACAGCGTGGCATTCTAGATTACTTTTCTTACTCTGAAACATGGCCTGTTT 1920
Qy 1921 gggagtgaggattcaagggtggtccaccgctccctactgcaaaaggcagtttaatt 1980
Db 1921 GGGAGTGGGGGATTCAAGGTGGTCCACCGCTGCCCTACTGCAAAATGGCAGTTTAAAT 1980
Qy 1981 ctatcttttggcttctgcagatggttgcgaattgatctcttaaccataatgaatgcagtcct 2040
Db 1981 CTTATCTTTGGCTTCTGCAGATGGTGCATTCATCTTAACCAATAATGGTCAGTCTCT 2040
Qy 2041 catctctgctcttcataggtgcacacttctgtgtttaagaagaggaaagctttgtacc 2100
Db 2041 CATCTCTGCTCTGCTTCTCATAGGTGGCCACTTGTGTGTTAAGAGAGGAAGCTTTGTACC 2100
Qy 2101 tttagagtgagtgaaatgaatgaatggcttggagtgagtgcaagacagcatatgattt 2160
Db 2101 TTTAGAGTGTAGTGAATGAATGAATGGCTGGAGTGCACCTGAGAACAGCATATGATTT 2160
Qy 2161 ctgcttgggggaaagaaatgatctatgaattggtgggtgtatttgagaagata 2220
Db 2161 CTTGCTTTGGGAAAGAAAGATGATGCTATGAATTTGGTGGTGTATTTGAGAAGATA 2220
Qy 2221 atcattgcttataatgaatgagctgaatttgataaaaccccaatacagctatgaagt 2280
Db 2221 ATCATTGCTTATGTCAATGAGCTGAATTTGATAAAACCCCAATAACAGCTATGAAGT 2280
Qy 2281 gctgggcaagtttactttttcttgatgttttctcacaact 2320
Db 2281 GCTGGCGCAGTTTACTTTTTTCTGATGTTCTTCACTCAACT 2320

RESULT 2
AX286764
LOCUS AX286764 2320 bp DNA linear PAT 21-NOV-2001
DEFINITION Sequence 3 from Patent WO0180896.
ACCESSION AX286764
VERSION AX286764.1 GI:17048799
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (sites)
AUTHORS Ratain,M.J., Innocenti,F. and Iyer,L.
TITLE Flavopiridol drug combinations and methods with reduced side
effects
JOURNAL Patent: WO 0180896-A 3 01-NOV-2001;
ARCH DEVELOPMENT CORPORATION (US)
FEATURES
source
1..2320
/organism="Homo sapiens"
/db_xref="taxon:9606"
19..1611
/note="unnamed protein product"
/codon_start=1
/protein_id="CAD12275.1"
/db_xref="GI:17048800"
/translation="MACITGWSPPLICVCLLLTCGFAEGAKLLVVPMDGSHWFTRSV
VEKILRGHEVYVMPVPSVOLGRSLNCTVKTYSITYLEDLDRFKFAFAHQAQV
RSIYSLMGSYNDIFDLFNSRSLFKDKLVEYLKESSEDAVLPDPFNCGLIVAKY
FSLPSVVFARGILCHYLEGAQCPAPLSYPRILLGSDAMTFKERVNIMHLEEH
LCHRFKNALIEASEILOTPTVEYDLYSHTSIWLRLDFVLDYKPKVMPNMFIGGIN
CHQKPLPMEFAYINASGEHIVFSLGSMVSEIPEKKAMATADALGKIPOTVLWRY
TGTPSNLANNITLVKLPONDLLGHMTRAFITHAGSHGVIESICNGVPVMPPLFG
DOMNARKMETKAGVTINLVLEMTSEDLNALKAVINDKSYKENIMRLSSLJHKDPVE
FDLNAVFWFEVVRHKGAPHLRPAADHLTWYQYHSLDVLGFLAVLVITFKCCA
IGYRKLCKGRVGRKHAHKSHTH"
BASE COUNT 608 a 495 c 547 g 670 t
ORIGIN
Query Match 100.0%; Score 2320; DB 6; Length 2320;

```

```

Best Local Similarity 100.0%; Pred. No. 0;
Matches 2320; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 ctcaactcacattctctgagcttgacagggtgagacagccccccttctctatgtg 60
Db 1 CTCACTGCAGTCTCTGATGAGCTTGACAGGGTGGACAGCCCTTCTCTATGTGTG 60
Qy 61 tgtctgctgacactgtggttccagaggcagggaagctactgtagtgcacattggtat 120
Db 61 TGTCTGCTGCTGACCTGTGGCTTTCGAGGCGAGGAAAGCTACTGGTAGTGCCATGGAT 120
Qy 121 gggagccactgtttccaccatgaggtcggtgtgagaaactcatctcaggggcatgag 180
Db 121 GGAGGACCTGTTTCAACATGAGGTGGTGGTGGAGAACTCATTTCTCAGGGGCATGAG 180
Qy 181 gtggtttagtcagtcagaggtgagttggcaactgggaagatcactgaattgcaactg 240
Db 181 GTGGTGTAGTCACTGCCAGAGGTGAGTTGGCACTGGGAGATCACTGAATGACAGTG 240
Qy 241 aaagacttattcaactctcatataccctggagatctggaccggagttcaaggtctttgcc 300
Db 241 AAGACTTATTCACCTTCATATACCTCTGGAGATCTGGACCGGGAGTTCAGGCTTTTGGC 300
Qy 301 catgtcctaaggaaagacaaagtagaagatatattctctattataatgggttcataaat 360
Db 301 CATGCTCAATGGAAGCACAAAGTACGAAGTATATATTTCTATTAATGGGTTCATAAAT 360
Qy 361 gacatttttgacttatttttttcaaatggcaggagttgttttaagagacaaaaattagta 420
Db 361 GACATTTTGTACTTATTTTTCACAAATTCAGAGGAGTTTGTAAAGACAAAAATTAGTA 420
Qy 421 gaactactaaaggagagatctttttagtcagtggtttctogactcttttgaactgtgac 480
Db 421 GAATACTTAAAGGAGAGATCTTTTATGATGAGTGTCTTCGATCCCTTTTGAATACTGTGC 480
Qy 481 ttaattgttgcaaaatatttctccctccctccgtggtctctccagggggaatactttgc 540
Db 481 TTAATTTGTGCAAAATATTTTCCCTCCCTCCGTGCTTCCGCCAGGGGAATACTTTGC 540
Qy 541 cactatcttgaagaagtgcaagtcctctctctctctctctctctctctctctctctctc 600
Db 541 CACTATCTTGAAGAAGGTGCAGAGTGCCTGCTCTCTCTTCTCTATGTGCCAGAAATCTC 600
Qy 601 ttagggttctcagatgccatgacttcaaggagagagtagcagaaacacatcatgcacttg 660
Db 601 TTAGGTTCTCAGATGCCATGACTTTCAAGGAGAGAGTAGCGAAACACATCATGCACATTG 660
Qy 661 gaggaacatttatccacgcttttttcaaaaatgccctagaaaatagcctctgaaatt 720
Db 661 GAGGAACATTATATGCGACCGTTTTCRAAATGCCCTAGAAATAGCCCTCTGAAAT 720
Qy 721 ctcaaacactgttagcggagtagtatctctacagccacacatcaatttggttgttcga 780
Db 721 CTCCAACACCTGTGTACGGAGTAGTATCTCTACAGCCACACATCAATTTGGTTGTTCGA 780
Qy 781 acggacttgggttggactatcccaaacccgtgagtcaccaacatgatctctcatgttggt 840
Db 781 ACGGACTTTGTTTGGACTATCCCAACCCGTGATGCCCAACATGATCTTCATTGGTGGT 840
Qy 841 atcaactgcatcagggaagcggttgctctatggaatttgaagcctacataatgtctct 900
Db 841 ATCAACTGCCATPCAGGGAAGCCGTTGGCTATGGAATTTGAAGCCCTACATTAATGCTTCT 900
Qy 901 ggaagaacatggaattgtggttttctctcttcttgggatacaatggtctcagaataatccagagaag 960
Db 901 GGAGAACATGGAATGTGGTTTTCTCTTTGGGATCAATGTGTCTCAGAAATTCAGAGAAG 960
Qy 961 aaagctatggcaattgctgatgttttgggcaaaaatccctcagacagctctgtgcggtac 1020
Db 961 AAAGCTATGGCAATGTGTGATGTCTTTGGGCAAAATTCCTCAGACAGTCTCTGTGGCGTAC 1020
Qy 1021 actggaaccccgaccatcgaattcttcgaaacacacgatactgttaagtgggtaccacaa 1080

```



```
/codon_start=1
/protein_id="CADI2274.1"
/db_xref="GI:17048798"
/translation="MAGTGTSPPLVCLLLTGGFAEAGKLLVVPMDGSHWETMSV
VEKLILRGHEVVVPMPSVWOLGSLNCTVKTSTVSTLEDLLDREFKAPAHQWKAQV
RSTYLLMGSYNDLFDLPFSNCRSLFKDKLYELKESSEFADVFLDPDFNGCLLYAKY
FSLPSVFAEGLCHYLEGAQCPAPLSYVPRILLGFSQAMTFKRYNHHIMHLEHL
LCHRFKNAIEAIEIQLQTPVEIYDLSYTSIMILLRTDFVLDYPKVPMNMFYGGIN
CHQGLPMEFEATYINASGHEGIVFSLGSMVSEIPEKKAMAIADALGKIPQTVLMRY
TGTSPSLNANTILVWLQNDLLGHPMTRAFITHAGSHGVYESTCNGVPMVWMLPFG
DQMDNAEMETKAGVTLNLEMTSEDLNALKAVINDKSYKENIMRLSLHKKDRPVE
PLDLAVFVVEYMRHKGAPHLRPAADHLTWYQVHSLDVLGFLAVLVLTVAFTFKCA
YGVKCLGKGRVKAHKSHT"
```

BASE COUNT 608 a 495 c 547 g 670 t

Query Match		100.00;	Score 2320;	DB 6;	Length 2320;
Best Local Similarity		100.00;	Pred. No. 0;		
Matches 2320;		Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	1	ctcagctgacgttctctgagctgtgacaggtgacagcccccctctctctatgtg	60		
Db	1	CTCAGCTGACGTCTCTGATGCTTGACAGGCTGGACAGCCCTCTCTATGTG	60		
Qy	61	tgctctgctgacgttctgcttccgagcgaggggaagctactgtagtgcctatgg	120		
Db	61	TGTC TGCTGCTGACCTTGCTGCTTCCGAGGCGACAGGAGCTACTGGTAGTGC	120		
Qy	121	ggagacactggttccaccatgaggtcggtggtgagaaactcattctcaggggcatag	180		
Db	121	GGGAGCCACTGGTTCACCATGAGGTGCGTGGTGGAGAACTCATCTCAGGGGGCAT	180		
Qy	181	gtggtgtagtcacgaggtgagttggcaactgggaagacacactgaattgcacagtg	240		
Db	181	GTGGTGTGTGTCATGCGCAGAGGTGAGTGGCAACTGGGAAGATCACTGAATTG	240		
Qy	241	agagcttattcaacttcataccctgaggtgaggtgagccggaggttcagggttttgc	300		
Db	241	AAGAATTATTCACCTTTCATATACCTCGAGGATCTGGACGGGAGTTCAAGGCT	300		
Qy	301	catgctcaatggaagacacagatcagaggtatataattctctattaaatgggttca	360		
Db	301	CATGCTCANTGGAAACACAGTACGAAGTATATATTCCTATTAATGGGTTCAT	360		
Qy	361	gacattttgacttatttttttcaaatgaggtggtttgttaaagacaaaattagta	420		
Db	361	GACATTTTGTACTTATTTTTCAAATGTCAGGAGTTTGTTTAAAGACAAAATTA	420		
Qy	421	gaataactaaagagaggtcttttgatgcaagttctctcgatcccttttgtaactgt	480		
Db	421	GAATACTTAAGGAGAGTTCCTTTGATGCAAGTCTTCGATCCCTTTTGATAACT	480		
Qy	481	ttaa-tgttgccaaatatttctccctccctccgtggtcttcgaggggaataacttgc	540		
Db	481	TTAAATGTTGCCAAATATTTCTCCCTCCCTCCGTGGTCTTCGCGAGGAATACT	540		
Qy	541	caactattcaagaaggtgcacagtgccctgctccctctctctctctctctctctc	600		
Db	541	CACATMCTTGAAGAGGTGCACAGTGGCCCTGCTCTCTCTCTATGTGCCAGAA	600		
Qy	601	ttaggttctcagatgccatgactttcaagagagagtagtcggaaccacatcatgca	660		
Db	601	TTAGGTTCTCAGATGCCATGACTTTCAGAGGAGAGTACGGAACCATCATGAC	660		
Qy	661	gagguacatttattgacacgttttttcaaaaatgccttagaataagctctctctct	720		
Db	661	GAGGACATTATTAATGCGCCCGTTTTCAAAATGCCCCTAGAAATAGCCTCTGA	720		
Qy	721	ctccnaacacactgtttacgagtagtatctctcagccacacatcaatttggttgcga	780		
Db	721	CTCCNAACACCTGTTACGAGTATGATCTCTACAGCCACACATCAATTTGTTGT	780		

```
Qy 781 acgagactgttttgactatcccaaacccggtgatcccaacatgatcttcattggt 840
|||||
Db 781 ACGGACTTGTGTTGACATATCCCAACCCGCGATGCCCACTGATCTCATGTGGT 840

Qy 841 atcaactgcacagggaaagcgttgcttatgaaatttgaagcctacataatgctt 900
|||||
Db 841 ATCAACTGCCATCAGGAAAGCCGTGCTTATGGAATTTGAAGCCTACATTAATGCTTCT 900

Qy 901 ggagacatggaattgtggtttctcttggatcaatgttctcagaaatccagagaag 960
|||||
Db 901 GGAGAACATGGAATGTGTTCTCTTTGGGATCAATGCTCTCAGAAATCCAGAGAAG 960

Qy 961 aaagtatgcaattctgctgctttgggcaaaatccctcagaacagtcctgtggcggtac 1020
|||||
Db 961 AAGGTATGCAATTTGCTGCTTTGGGCAAAATCCCTCAGACAGTCTCTGTGGCGGTAC 1020

Qy 1021 actggaaccgcacatcgaaatcttgcaaacacacatacttgaagtggctaccacca 1080
|||||
Db 1021 ACTGGAACCGACCATCGAATCTTGCGAACAACACGATACTTGTTAAGTGGCTACCCCAA 1080

Qy 1081 aaactctgttgctcacccatgacccgtgaccttatccaccatgctggttccatggt 1140
|||||
Db 1081 AACGATCTGCTTGTGCTCCCGCATGACCCGCTGCTTATCACCATGCTGGTCTCCATGCT 1140

Qy 1141 gtttatgaaagcatatgcaatggcggttcccatggtgatgacccctgttggtagacag 1200
|||||
Db 1141 GTTTATGAAGCATATGCAATGGCGTTCCCATGCTGATGATGTCCTTGTGTTGTGATCAG 1200

Qy 1201 atgacaactcaaaagcagactgaggaactaaaggagctggagtgacctgaattgtgaa 1260
|||||
Db 1201 ATGGACAATGCAAGCGCATGGAGACTAAGGGAGCTGGAGTGAACCTGGAATGTTCTGGA 1260

Qy 1261 atgactctgaaatttagaaatgctctaaaagcagtcataatgacaaaagtacaaag 1320
|||||
Db 1261 ATGACTTCTGAAGATTGAAATGCTCTAAAAGCAGTCAATCAATGACAAAAGTTACAAG 1320

Qy 1321 gagaacatcatgctgctctccagcttcacaagaccgcccgggtggagcgtggaactg 1380
|||||
Db 1321 GAGAACATCATGCGCTCTCCAGCCTTCACAAGACC GCCCGGTGGAGCGCTGGACCTG 1380

Qy 1381 gccgtgttcgggtgaggttctgatgaggcaacaaggcgccacacacctgcccgcga 1440
|||||
Db 1381 GCCGTGTTCTGGTGGAGTGTGATGAGGCACAAGGGCGGCCACACCTGCGCCCGCCA 1440

Qy 1441 gccacacacctcacctggttaccagttaccattctctgagcgtgattgttctctggcc 1500
|||||
Db 1441 GCCCACACCTCACCTGGTACCAGTACCATTCTCTGGAGCTGATGGTTCTCTTGGCC 1500

Qy 1501 gtcgtgctgacagtgccctctacacctttaaatgtgtgttatggtgacctgacgaaatgc 1560
|||||
Db 1501 GTCGTGTGACAGTGGCCCTTCATCACCCTTTAAATGTTGTGCTTATGCTACCGGAAATGC 1560

Qy 1561 ttggggaaaaagggcgagttlaaagaagcccaaaatccaagacccttgagaagtgggt 1620
|||||
Db 1561 TTGGGGAAAAAGGGCGAGTTAAGAAAGCCCAACAAATCCAAGACCCATTTGAGAAATGGGT 1620

Qy 1621 gggaaataaggttaaaatttgaacattcccttagtcaattcccaacttgaacacagaatc 1680
|||||
Db 1621 GGGAAATAAGGTAAAAATTTGAACCATTCCTTAGTCAATTTCCAAACTTGAACACAGAAATC 1680

Qy 1681 agtgttaattcaattttcttattagaagaactcttgataaaataatcagccccag 1740
|||||
Db 1681 AGTGTAAAATTCATTTTATTTATTAAGGAATACTTTGTCATAAATTAATCATAGCCCCAG 1740

Qy 1741 agtctttaaataatctcttaaataaaaaataatagactcgctagtcagtaaatattt 1800
|||||
Db 1741 AGTCTTTAAAAAATCTCTTAAAAATAAAAAATAATAGACTCGCTAGTCACTAGTAAAGATATT 1800

Qy 1801 gaatatatcgccccctcgctgtcttttaacagatcaacatgacgtgcaattttcaga 1860
|||||
Db 1801 GAATATGTAICGTGCCCCCTCGGTGCTTTTGTGATCAGGATGACATGTCCTATTTTTCAGA 1860

Qy 1861 ggcgtgcajacaggtgctggaattcttagattacttttttacttgaaacatggcctggtt 1920
```

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 21, 2002, 11:57:31 ; Search time 8248.67 Seconds
(without alignments)
5885.745 Million cell updates/sec

Title: US-09-835-082-3

Perfect score: 2320

Sequence: 1 ctacagtcagttcttgat.....ttctgatttctactaacat 2320

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1797656 seqs, 10463268293 residues

Total number of hits satisfying chosen parameters: 3595312

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl.*

- 1: gb_ba.*
- 2: gb_htg.*
- 3: gb_in.*
- 4: gb_on.*
- 5: gb_ov.*
- 6: gb_pat.*
- 7: gb_ph.*
- 8: gb_pl.*
- 9: gb_pr.*
- 10: gb_ro.*
- 11: gb_sts.*
- 12: gb_sy.*
- 13: gb_un.*
- 14: gb_vl.*
- 15: em_ba.*
- 16: em_fun.*
- 17: em_hum.*
- 18: em_in.*
- 19: em_mu.*
- 20: em_on.*
- 21: em_or.*
- 22: em_ov.*
- 23: em_pat.*
- 24: em_ph.*
- 25: em_pl.*
- 26: em_ro.*
- 27: em_sts.*
- 28: em_un.*
- 29: em_vl.*
- 30: em_htg_hum.*
- 31: em_htg_inv.*
- 32: em_htg_other.*
- 33: em_htgo_inv.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

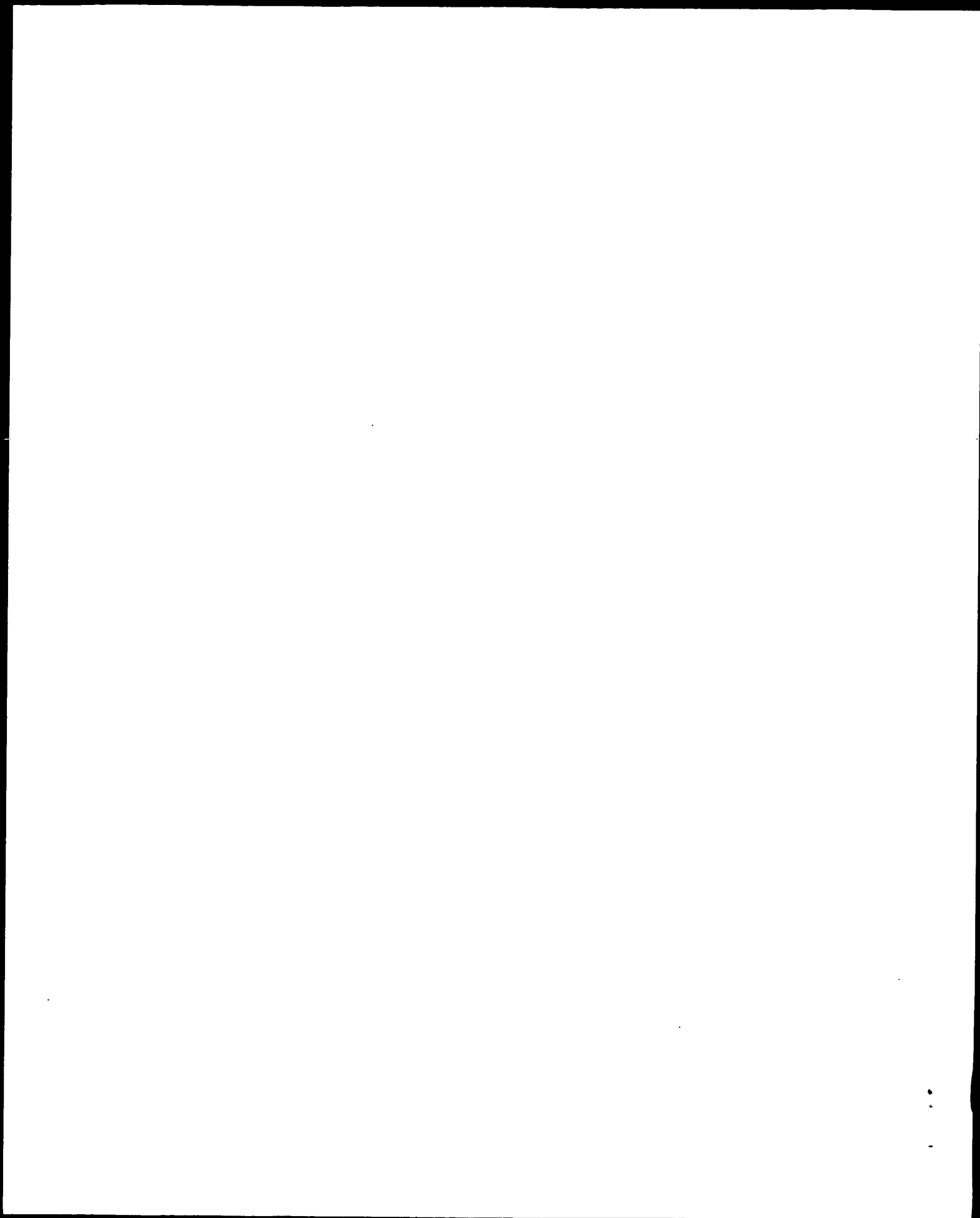
SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
------------	-------	-------	--------	----	----	-------------

RESULT 1
AX286762
LOCUS AX286762 2320 bp DNA linear PAT 21-NOV-2001
DEFINITION Sequence 1 from Patent WO0180896.
ACCESSION AX286762
VERSION AX286762.1 GI:17048797
KEYWORDS human.
SOURCE Homo sapiens
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (sites)
AUTHORS Ratain,M.J., Innocenti,F. and Iyer,L.
TITLE Flapoviridol drug combinations and methods with reduced side effects
JOURNAL Patent: WO 0180896-A 1 01-NOV-2001;
FEATURES ARCH DEVELOPMENT CORPORATION (US)
source Location/Qualifiers
1..2320 /organism="Homo sapiens"
/db_xref="taxon:9606"
19..1611 /note="unnamed protein product"

ALIGNMENTS

1	2320	100.0	2320	6	AX286762	AX286762	Sequence 1 from Patent WO0180896.
2	2320	100.0	2320	6	AX286764	AX286764	Sequence
3	2320	100.0	2320	9	AF056188	AF056188	Homo sapi
4	2218.4	95.6	2363	9	AF462267	AF462267	Homo sapi
5	2207.4	95.1	2394	9	AF462268	AF462268	Homo sapi
6	2194.6	94.6	2423	9	BC020971	BC020971	Homo sapi
7	2156.8	93.0	2392	9	S55985	S55985	Homo sapien
8	2076	89.5	2396	9	AF104336	AF104336	Macaca fa
9	1895.8	81.7	2003	9	AK025403	AK025403	Homo sapi
10	1636.6	70.5	2368	9	H0MHUGB2	H0MHUGB2	Human billr
11	1629	70.2	2388	9	BC019861	BC019861	Homo sapi
12	1610.8	69.4	2351	9	H0MHUGB1	H0MHUGB1	Human billr
13	1539.2	66.3	2422	6	AX336711	AX336711	Sequence
14	1539.2	66.3	2422	6	AX336711	AX336711	Sequence
15	1493.8	64.4	1593	9	AF030310	AF030310	Homo sapien
16	1484.2	64.0	1593	9	HS089507	HS089507	Human UDP-g
17	1477	63.7	1621	9	AF104338	AF104338	Macaca fa
18	1470.6	63.4	2440	9	AF104337	AF104337	Macaca fa
19	1463.6	63.1	1593	9	HS089508	HS089508	Human UDP-g
20	1157.8	49.9	2241	10	AF461735	AF461735	Rattus no
21	1129.4	48.7	2203	10	BC019434	BC019434	Mus muscu
22	1123.4	48.4	1590	4	AF011902	AF011902	Oryctolag
23	1110.4	47.9	2155	10	MUSUGTP	MUSUGTP	Mus musculu
24	1105.2	47.6	2332	4	AB018478	AB018478	Ovis arie
25	1105	47.6	2282	10	AF461738	AF461738	Rattus no
26	1004.4	43.3	1561	9	H0MUGT03	H0MUGT03	Human UDP-g
27	1004.4	43.3	68770	9	AC006985	AC006985	Homo sapi
28	1004.4	43.3	176619	2	AC026497	AC026497	Homo sapi
29	1004.4	43.3	198872	9	AF297093	AF297093	Homo sapi
30	937	40.4	2301	10	RNU75903	RNU75903	Rattus norv
31	878.6	37.9	1255	9	AF297091	AF297091	Homo sapi
32	878.6	37.9	148438	9	AC019072	AC019072	Homo sapi
33	873	37.6	1602	9	AF360121	AF360121	Macaca mu
34	872.2	37.6	1647	9	AF104339	AF104339	Macaca fa
35	868.2	37.4	1602	9	AY029169	AY029169	Macaca mu
36	819.8	35.3	2816	4	AF039138	AF039138	Felis cat
37	799.8	34.5	183826	2	AC091778	AC091778	Papio cyn
38	794	34.2	2300	10	MMU16818	MMU16818	Mus musculu
39	788.4	34.0	2037	4	OCU09101	OCU09101	Oryctolagus
40	783.2	33.8	1765	10	RATUDPGTA	RATUDPGTA	Rat billirub
41	781.8	33.7	2343	10	AF461737	AF461737	Rattus no
42	777.9	33.6	1244	9	HS042604	HS042604	Human UDP-g
43	777.8	33.5	1927	10	RATUDPGT	RATUDPGT	Rat UDP-glu
44	774	33.4	2239	10	RNU20551	RNU20551	Rattus norv
45	772.2	33.3	2484	10	RATUGT	RATUGT	Rat 3-methy



QY 1459 taccagtaccattccttggacgtgattggtttcctcttggccgtgctgacagtggcc 1518
|||||
Db 1441 TACCAGTACCAATTCCTTGGACGTGATTGGTTTCCTTTGGCCGTCGTCGACAGTGGCC 1500
QY 1519 ttcatcaccttttaaatgttgtcttatgctaccggaatgcttgggaaaaaaggcgca 1578
|||||
Db 1501 TTATCACCTTTAAATGTGTGTGTATGGCTACCGGAATGCTTGGGGNAAAAAGGCGCA 1560
QY 1579 gtttaagaagcccaacaatccaagaccattga 1611
|||||
Db 1561 GTTAAGAAAGCCCAAAATCCAAAGACCCATTGA 1593

Search completed: August 21, 2002, 11:57:31
Job time: 32834 sec

QY 1805 atgtatcgtgccccctccgtgtctttgatcagatgacatgtgcatttttccaggac 1864
 |||||
 Db 1811 ATGATATCGTGGCCCTCCGCTGCTTTTGTATCAGGATGACATGTGCCATTTTTCAGAGGAC 1870
 |||||
 QY 1865 gtccagacagcgtggcattctagattacttttttactctgaataatgctgtttggga 1924
 |||||
 Db 1871 GTGCAGACAGGCTGCACTCTAGATTACTTCTCTCTGAAACATGCGCTGTTGGGA 1930
 |||||
 QY 1925 gtgaggattcaaaagggtggccaccgctgccccctactgccaatggcgattttaattta 1984
 |||||
 Db 1931 GTGCGGATTCAAAGGTGGTCCCAACGCTGCCCTACTGCAAAATGCGAGTTTAACTTAA 1990
 |||||
 QY 1985 tcttttgcttcagatggttgcaattgactcttaacaaatgactgactgactcactc 2044
 |||||
 Db 1991 TCTTTTGGCTCTCGCAGATGTTGCAATTTGATCTTAAACCAATATGTCAGTCTCTCATC 2050
 |||||
 QY 2045 tctgctccttcctagtgccaccctgtgtgttttaagaagggaagctttgtacottta 2104
 |||||
 Db 2051 TCTGTCTCTCTCATAGTGCACCTGTGTGTGTTTAAAGAAAGGAAGCTTTGTACCTTAA 2110
 |||||
 QY 2105 gaagttagtgaatgaatgaatgctgtgagtgacactgagacagcatatgattcttg 2164
 |||||
 Db 2111 GAGTGTAGTGGAATGAATGAATGGCTTGGAGTGCACTGAGACACATATGATTCTTG 2170
 |||||
 QY 2165 ctttggggaaaaaagaatgactgctatgaaatgggtgggtgtattttgagaagataatca 2224
 |||||
 Db 2171 CTTTTGGGAAAAAAGAAATGATGCTATGAAATTTGGTGGTGTATTTTGAAGAAGATAATCA 2230
 |||||
 QY 2225 ttcttatctcaaatgagctgaatttgataaaaccccaataacacgctatgaagtctg 2284
 |||||
 Db 2231 TTGCTTATGTCAATTTGGAGCTGAATTTGATATAAACCCCAAAATACAGCTATGAAGTCTG 2290
 |||||
 QY 2285 ggcaagttacttttttctgactgtttctcacaact 2320
 |||||
 Db 2291 GGCAAGTTTACTTTTTTCTGATGTTTCTTCAACT 2326
 |||||

RESULT 13
 LOCUS AX336711 2422 bp DNA linear PAT 09-JAN-2002
 DEFINITION Sequence 7220 from Patent WO0194629.
 ACCESSION AX336711
 VERSION AX336711.1 GI:18127430
 KEYWORDS
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (sites)
 AUTHORS Young,P.E., Augustus,M., Carter,K.C., Ebner,R., Endress,G.,
 Horrigan,S., Soppet,D.R. and Weaver,Z.
 TITLE Cancer gene determination and therapeutic screening using signature
 gene sets
 JOURNAL Patent: WO 0194629-A 7220 13-DEC-2001;
 Avalon Pharmaceuticals (US)
 FEATURES
 source location/qualifiers
 1..2422
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 BASE COUNT 653 a 516 c 555 g 698 t
 ORIGIN

Query Match 66.3%; Score 1539.2; DB 6; Length 2422;
 Best Local Similarity 82.0%; Pred. No. 0;
 Matches 1870; Conservative 0; Mismatches 398; Indels 12; Gaps 8;
 QY 47 ttctctatgtgtgtgtgctgctgactgtggttttgcaggcagggaagctactgtg 106
 |||||
 Db 119 TTTCTGACAGGGGTTTCTTCTTAGCTTTTGGGGCATGGTTGTAGGTGACAAAGCTGCTGG 178
 |||||
 QY 107 tagtgcccatggatggagccactggttcaccatgagtgctggtgggagaaactcattc 166
 |||||

Db 179 TGTGTCCTCAGAGCGGAAGCCACTTTTGTATGAAAGATATAGTTGAGGTTCTCAGTG 238
 |||||
 QY 167 tcaggggcatagagtggtgttagtcatgccagaggtgagttggcaactgggaagatcac 226
 |||||
 Db 239 ACCGGGTCATGAGATTGATGGTGGTGCCTGAAGTTAATTTGCTTTTGAAGAATAACA 298
 |||||
 QY 227 tgaattgcacagtgaaagacttatacaacttcataaccctggagatctggaccggaggt 286
 |||||
 Db 299 AATATCTACACAAAGAAAATCTATCCAGTGCCTGATGACCAAGAAAGAGCTGAAGAACCGTT 358
 |||||
 QY 287 tcaaggcttttgcctatgctcgaatggaaagcacaagtcagaagtatattctctattaa 346
 |||||
 Db 359 ACCAATCATTTTGGAAACAATCACTTTGTGTAGCGATCATTTCTTAACCTGCTCTCAGACAG 418
 |||||
 QY 347 tgggttcatacaatgaca---tttttgacttatttttttcaaatgacagaggtttgtta 403
 |||||
 Db 419 AGTACAGGAATAACATGATTGTTTATTTGGCTGTACTTCATCACTGCCAGAGCCTCTCGC 478
 |||||
 QY 404 aagacaaaaatttagtagaataacttaaaggagagttctttttgatgcagtggtttctgatc 463
 |||||
 Db 479 AGCAGAGGACACACCTCAACTCTTTTAAAGGAGAGCAAGTTTGTATGCTCTTTTTCAGAGCC 538
 |||||
 QY 464 cttttgataaactgtggtcttaattgttgcaaatattttccctccctccctccctgtcttg 523
 |||||
 Db 539 CAGCCTTACCTGTGGGGTGTCTCTGGCTGAGTATTTTGGCCCTTACATCTGTGTACTCT 598
 |||||
 QY 524 ccagggaataactttggcactatcttgaagaaggtgcacagtgccctgctctctctctct 583
 |||||
 Db 599 TCAGGGGTTTTCCGTGTTCCCTGGAGCATACATTTTCAGAGAGCCAGACCTCTGTCTCT 658
 |||||
 QY 584 atgtcccaagaattctttaggttctcagatgcatgcttcaagagagagagacgga 643
 |||||
 Db 659 ACATTCACAGGTGTACACAAAGTTTTCAGACCATGACTTTTCCCAACGAGTGCCCA 718
 |||||
 QY 644 accacatcatgcacttgaggaaacatttattatgccaccgttttttccaaaatgccctag 703
 |||||
 Db 719 ACTTCCTTTGTAATTTGTGGAGCCCTATCTATTTTATTTGCTGTGTTTCAAGATATGAGA 778
 |||||
 QY 704 aaatacctctgaattctccaaacactgttcacggagatgatctctcacagccacacat 763
 |||||
 Db 779 AACTCCCATCAGCTGCTCTCAAGAGAGATGT---GGATATAATCACTTATCAGAGGTCT 835
 |||||
 QY 764 caatttggtgtgcagacggaactttgttttgactatccccaaacccgtgagtcacca 823
 |||||
 Db 836 CTGTTTGGCTGTTAAGATATGACTTTGTGCTTGAATATCTAGCCGGTCAATGCCCAACA 895
 |||||
 QY 824 tgatctcatttggtgtatcaactgcatcagggaaagccgttgctatggaatttgaag 883
 |||||
 Db 896 TGGTCTTCATTGGAGGTATCACTGTAAAGAGAGAAAGACTTGTCTCAGGAATTTGAAG 955
 |||||
 QY 884 cctacattatgctcttgaggacatggaattgtggttttctctctttgggatacaatggtct 943
 |||||
 Db 956 CTACATTAATCTCTCTGGAGAACATGGAATTTGGTGTGTTTCTTTGGGATCAATGGTCT 1015
 |||||
 QY 944 cagaaattccagagaagaagctatggcaattgtgtgcttttgggcaaaaacctcaga 1003
 |||||
 Db 1016 CAGAAATTCAGAGAAGAAAGCTATGGCAATTTGCTGATGCTTTTGGCAAAAACCTCAGA 1075
 |||||
 QY 1004 cagtctgtggtggtgacacggaccccgaccatcgaaatcttctggaacacacagatacttg 1063
 |||||
 Db 1076 CAGTCTGTGGCGGTACACTTGAACCCGACCATCGAATCTTTGGGAAACACAGGATACTTG 1135
 |||||
 QY 1064 ttaagtggctaccccaaaacgactctgtgttcacccgagacccgtgcttttaccacc 1123
 |||||
 Db 1136 TTAAGTGGCTACCCCAAAACGATCTGTTGGTCAACCCGATGACCCGTGCTTTTATCACCC 1195
 |||||
 QY 1124 atgctggttccactggtgtttatgaagcatatgcaatggcggttccatggtgatgac 1183
 |||||
 Db 1196 ATGCTGTTCCCATGGTGTATTAAGAGCATATGCAATGGCGTTCCTCATGGTGTATGTC 1255
 |||||
 QY 1184 ccttggttggtgatcagatggacaatgcaaacgcatggagactaaggagcgtggagtg 1243
 |||||

10

Db	726	AACTGCGCATCAGCTGTCTCTCAAGAGAGATGTGGATATAATCACCTTATATATCAGAAGGTCT	785
Qy	764	caatttggtttgttcgaagcgactttgttttgactatcccaaacccggtatgcccaca	823
Db	786	CTGTTTGGCTGTTAAGATATGACTTGTGCTGAATACCTTAGCGCGGTATGCCACACA	845
Qy	824	tgatcttcattggttggttatcaactgcoactcagggaagccgtgtcctatggaaatttgaag	883
Db	846	TGGTCTTTCATTGGAGGTATCAACTCTAAGAGAGAGGAAGACATTTGCTCTCAGGAATTTGAAG	905
Qy	884	ccatcataaactctcttgagaaatgaatttgttttctcttggaaatcaatggctct	943
Db	906	CCTACATTAATGCTCTTGGAGAACATGAATGTGTGTTTTCTTTGGGATCAATGTGCT	965
Qy	944	cagaaattccagagaagaaagctaaggcaattgctgatgcttttggcgaataaccctcaga	1003
Db	966	CAGAAATTCACAGAGAAAGACTATGCGCACTGCTGATGCTTTGGCGCAAAATCCCTCAGA	1025
Qy	1004	caftcctgtgcagtaacatgaaacccagacatcgaactctgcaactctgcgaacaacagtaacttg	1063
Db	1026	CAGTCTGTGGCGGTACTCGAAGCCGACCATCGAAHCTTGGGAAACACACAGACTCTTG	1085
Qy	1064	ttaagtggtaccaccaacgatctgttggtcacccgatgacccgtgaccttatacacc	1123
Db	1086	TTAAGTGGCTACCCCAAAACGATCTGCTGTGTCAACCCGATGACCCGTGCTTTATCACACC	1145
Qy	1124	atgctggttccccatggttttatgaaagcatatgcgaatggcggttccccatggtgatgctgc	1183
Db	1146	ATGCTGGTTCCTCCATGGTGTATTAGAAACATATGCAATGGCGTTCCCATGGTGATGATGC	1205
Qy	1184	ccttgtttggtatcagatggacaatgcnaagcgcatggagactaaaggagagctggagtga	1243
Db	1206	CCTTGTTGGTATCATGATGGACAATGCAAGAGCGCATGGAGACTAAGGGAGTGGAGTGA	1265
Qy	1244	ccctgaaattcttggaatgaactctctgaaatttagaaatgctctctaaagcagtcata	1303
Db	1266	CCCTGGAATGCTCGAAATGACTTCTGAAGATTTAGAAATGCTCTAAAAGCAGTCATCA	1325
Qy	1304	atgacaaaagttaaaaggaaacatactgcgcctctccagcttcaagagctcaagagccgcctcg	1363
Db	1326	ATGACAAAAGTTACAAAGGAAACAATCATGCCCTCTCCAGCCTTCCAGAGACCCGCCGG	1385
Qy	1364	tggagccgttggaacctgcccgtgttctgggtggagtgttgatagggcacaaaggcgcg	1423
Db	1386	TGGAGCCGCTGGACCTGGCCGCTTCTGGGTGGAGTTTGTGATGAGCACAAAGGCGCGC	1445
Qy	1424	caactctgcccgcagcccaagacctcaactgtgtaccagtaccattctctggagctga	1483
Db	1446	CACACCTTGCCCCCGACCCACGACCTTACCTGGTACCGATCCATCTCCTTTGGACGTGA	1505
Qy	1484	ttggtttctcttggccgtgctgcagtgagtgcccttcatcactttaaatgtgtgctt	1543
Db	1506	TTGGTTTCCCTTGGCCGCTGCTGACAGTGGCCTTTCATACCTTTTAAATGTTGTGGTT	1565
Qy	1544	atgctaccggaatgcttgggaaaaaaggcgagtaagaagaccacaaatccaaga	1603
Db	1566	ATGCTACC GGAAATGCTTGGGAAANAAGGCGGTGTAGNAAGCCCAAAATCCANAGA	1625
Qy	1604	cccattgagaagtgggtggaaaaataaggttaaaatttgaccatttccctagtcatctcca	1663
Db	1626	CCCATTTGAGAAGTGGGTGGGAAATAAGGTAAATTTGAACCATTTCCCTAGTCATTTCGA	1685
Qy	1664	aacttgaacaggaatcagtggttaaatctatttattctttaaggaataactttgat	1723
Db	1686	AACTTGAACAGATCAGTGTAAATTCATTTATTCATTATTAAGGAATACTTTGCAT	1745
Qy	1724	aaattaatagccccagatgctttaaaaaattctcttaataaaaaataatagactcgct	1783
Db	1746	AAATTAATACGCCACAGAGTCTTTAAAAAATCTCTTTAAATAAAAATAATAGACTTCGT	1805
Qy	1784	agtcatgaagatatgaatatgatatggtccccctcggtgtcttttgatcaggatgac	1843

D	b	1806	AGTCAGTAAGAATATTGAATATGTATCTGCGCCCTCCTGGTGCTTCTTGATCAGATGAC	1865
Q	y	1844	atgtgcatttttcagaggacctgcagacaggctggcattctagatacttcttctact	1903
D	b	1866	ATGTGCATTTTTTCAGAGGACGTGCACAGCGTGCATTTCTAGATTACTTTCTTACTC	1925
Q	y	1904	tgaacaatgacctgttttggaagtgcgggattcaaaaggtagccaccgcctgccctactg	1963
D	b	1926	TGAACAATGGCCTGTGTGGAGGTGGGGATTCAAAGGTGGTCCCACGCTGCCCTTACTG	1985
Q	y	1964	caaatggcgagtttaactcttcttcttggctctgcagatggttccaattgatccttaac	2023
D	b	1986	CAAATGSCAGTTTAAATCTTATCTTTTGGCTTCTGCAGATGTTGCAATTGATCCTTAAC	2045
Q	y	2024	caataatgtcagtcctcatctctctctcgtctcataggtcacaccttgtgtttaag	2083
D	b	2046	CAATAATGTGTGAGTCCATCTCTCTCTCTCATAGTGCCACCCTGTGTGTTTTAAG	2105
Q	y	2084	aagggaagcttcttacctcttagagtagtgtaaagaatgaatggcttggagtgcactg	2143
D	b	2106	AAGGGAAGCTTTGTACCTTTTAGAGTGTAGGTGAATGAATGAATGGCTGGAGTGCACTG	2165
Q	y	2144	agaacagcatgattcttctgtcttgggaaaaaaaagaatgactatgaattgdtgggtg	2203
D	b	2166	AGAACAGCATATGATTCTCTCTTTGGGAAAAGAAGATGCTATGAATTTGTTGGGTG	2225
Q	y	2204	gtgtatttgagaagaataatcattgcttatgtaaaatggagctgaattgataaaaccca	2263
D	b	2226	GTGTATTTCAGAGAATAATCATTTGCTTATGTCAAATGGAGCTGAATTTGATAAAAACCCA	2285
Q	y	2264	aatcacagcatgaagtgcctggggcagtttacttttttctctgatgttctctacaact	2320
D	b	2286	AAATACAGCTATGAAGTCTTGCGCAAGTTTACTTTTTTCTGTGATTTCTCTACAACT	2342
RESULT	12			
HUMHUGBR1				
LOCUS				
DEFINITION				
ACCESSION				
VERSION				
KEYWORDS				
SOURCE				
ORGANISM				
REFERENCE				
AUTHORS				
TITLE				
JOURNAL				
MEDLINE				
FEATURES				
gene				
CDS				

LAKYLSIPAVFWRYICDLDPKGQCNPSSYIKLLTNTSDHMTFLQRVKNULPL
ALSTICHTSPAYLASELFEVRSVLDVSLVSWLGEFGRVMDYPRIPMPNMYFI
GTCNACPKSQFEAFYINASEGIVYFVSLGVSFPEKKAMATADALGIPOTV
LWRYGTNPANLANTIKVYLPONDLSGHPMTARTITHAGSHGVYESICNGPYMFW
PFGYDOMONAKMETGAGVYLPNLEMTSEDLNALKYINDKSYENKRLNLSLKH
RVEPDLADLAFVFEWRYMRKHGKSPHRAAHDLTWYQVSHLVDYFLPAAVLTVF
KCCAYSYRKLGLGKGRVYKHAASPHLTH

siq_peptide

```

03: 1.25
/gene="UGT2B"

```

/gene="UGT2B"

```
/note="G00-127-753"
```

96. .1631

0800-766-9333

```

/gene=UG12D
/EC_number="2.4.1.17"

```

/note="G00-127-753"

```
/product="UDP-glucose"
```

В" продукт-оф гуадиомосутилатерасе з тамтлү полимерасе

534 c 573 a 657 t

3
5
7
9
11
13
15
17
19
21
23
25
27
29
31
33
35
37
39
41
43
45
47
49
51
53
55
57
59
61
63
65
67
69
71
73
75
77
79
81
83
85
87
89
91
93
95
97
99
101
103
105
107
109
111
113
115
117
119
121
123
125
127
129
131
133
135
137
139
141
143
145
147
149
151
153
155
157
159
161
163
165
167
169
171
173
175
177
179
181
183
185
187
189
191
193
195
197
199
201
203
205
207
209
211
213
215
217
219
221
223
225
227
229
231
233
235
237
239
241
243
245
247
249
251
253
255
257
259
261
263
265
267
269
271
273
275
277
279
281
283
285
287
289
291
293
295
297
299
301
303
305
307
309
311
313
315
317
319
321
323
325
327
329
331
333
335
337
339
341
343
345
347
349
351
353
355
357
359
361
363
365
367
369
371
373
375
377
379
381
383
385
387
389
391
393
395
397
399
401
403
405
407
409
411
413
415
417
419
421
423
425
427
429
431
433
435
437
439
441
443
445
447
449
451
453
455
457
459
461
463
465
467
469
471
473
475
477
479
481
483
485
487
489
491
493
495
497
499
501
503
505
507
509
511
513
515
517
519
521
523
525
527
529
531
533
535
537
539
541
543
545
547
549
551
553
555
557
559
561
563
565
567
569
571
573
575
577
579
581
583
585
587
589
591
593
595
597
599
601
603
605
607
609
611
613
615
617
619
621
623
625
627
629
631
633
635
637
639
641
643
645
647
649
651
653
655
657
659
661
663
665
667
669
671
673
675
677
679
681
683
685
687
689
691
693
695
697
699
701
703
705
707
709
711
713
715
717
719
721
723
725
727
729
731
733
735
737
739
741
743
745
747
749
751
753
755
757
759
761
763
765
767
769
771
773
775
777
779
781
783
785
787
789
791
793
795
797
799
801
803
805
807
809
811
813
815
817
819
821
823
825
827
829
831
833
835
837
839
841
843
845
847
849
851
853
855
857
859
861
863
865
867
869
871
873
875
877
879
881
883
885
887
889
891
893
895
897
899
901
903
905
907
909
911
913
915
917
919
921
923
925
927
929
931
933
935
937
939
941
943
945
947
949
951
953
955
957
959
961
963
965
967
969
971
973
975
977
979
981
983
985
987
989
991
993
995
997
999

NOTES

Query Match	70 5%	Score	1636	6:	DB	9:	Length	2368:
-------------	-------	-------	------	----	----	----	--------	-------

Query Match	70.5%;	Score 1636.6
Best Local Similarity	83.4%;	pred NC 0:

Best Local Similarity 83.4%; PRED. NO. 0;
Matches 1871: Conservative 0; Mismatches 360; Indels 2; Cans 1.

[illegible][illegible]

db 101

150

DD TGT C TGGGC T GAGAG T GGAAGG T GTTGG TGGTGCAGCCCTGGCTTCAGCAT T60

[illegible]

QY 141 GAGGCGGCCTGGGGAACCTCATCTCAGGGGGCATGAGGTGGTCTGTAGTCATGCCAGA 200

161

DB 161 GCGGGAGGCCCTTGCGGAGCTCCATGCCAGAGGCCACCAGGCGGTGGTCCCTCACCCCGA 220

[illegible]

QY 201 ggttggtcggcaactgggaagatcactgaattgcacagtgaagacttattcaactcata 260

26
 27
 28
 29
 30
 31
 32
 33
 34
 35
 36
 37
 38
 39
 40
 41
 42
 43
 44
 45
 46
 47
 48
 49
 50
 51
 52
 53
 54
 55
 56
 57
 58
 59
 60
 61
 62
 63
 64
 65
 66
 67
 68
 69
 70
 71
 72
 73
 74
 75
 76
 77
 78
 79
 80
 81
 82
 83
 84
 85
 86
 87
 88
 89
 90
 91
 92
 93
 94
 95
 96
 97
 98
 99
 100
 101
 102
 103
 104
 105
 106
 107
 108
 109
 110
 111
 112
 113
 114
 115
 116
 117
 118
 119
 120
 121
 122
 123
 124
 125
 126
 127
 128
 129
 130
 131
 132
 133
 134
 135
 136
 137
 138
 139
 140
 141
 142
 143
 144
 145
 146
 147
 148
 149
 150
 151
 152
 153
 154
 155
 156
 157
 158
 159
 160
 161
 162
 163
 164
 165
 166
 167
 168
 169
 170
 171
 172
 173
 174
 175
 176
 177
 178
 179
 180
 181
 182
 183
 184
 185
 186
 187
 188
 189
 190
 191
 192
 193
 194
 195
 196
 197
 198
 199
 200
 201
 202
 203
 204
 205
 206
 207
 208
 209
 210
 211
 212
 213
 214
 215
 216
 217
 218
 219
 220
 221
 222
 223
 224
 225
 226
 227
 228
 229
 230
 231
 232
 233
 234
 235
 236
 237
 238
 239
 240
 241
 242
 243
 244
 245
 246
 247
 248
 249
 250
 251
 252
 253
 254
 255
 256
 257
 258
 259
 260
 261
 262
 263
 264
 265
 266
 267
 268
 269
 270
 271
 272
 273
 274
 275
 276
 277
 278
 279
 280
 281
 282
 283
 284
 285
 286
 287
 288
 289
 290
 291
 292
 293
 294
 295
 296
 297
 298
 299
 300
 301
 302
 303
 304
 305
 306
 307
 308
 309
 310
 311
 312
 313
 314
 315
 316
 317
 318
 319
 320
 321
 322
 323
 324
 325
 326
 327
 328
 329
 330
 331
 332
 333
 334
 335
 336
 337
 338
 339
 340
 341
 342
 343
 344
 345
 346
 347
 348
 349
 350
 351
 352
 353
 354
 355
 356
 357
 358
 359
 360
 361
 362
 363
 364
 365
 366
 367
 368
 369
 370
 371
 372
 373
 374
 375
 376
 377
 378
 379
 380
 381
 382
 383
 384
 385
 386
 387
 388
 389
 390
 391
 392
 393
 394
 395
 396
 397
 398
 399
 400
 401
 402
 403
 404
 405
 406
 407
 408
 409
 410
 411
 412
 413
 414
 415
 416
 417
 418
 419
 420
 421
 422
 423
 424
 425
 426
 427
 428
 429
 430
 431
 432
 433
 434
 435
 436
 437
 438
 439
 440
 441
 442
 443
 444
 445
 446
 447
 448
 449
 450
 451
 452
 453
 454
 455
 456
 457
 458
 459
 460
 461
 462
 463
 464
 465
 466
 467
 468
 469
 470
 471
 472
 473
 474
 475
 476
 477
 478
 479
 480
 481
 482
 483
 484
 485
 486
 487
 488
 489
 490
 491
 492
 493
 494
 495
 496
 497
 498
 499
 500
 501
 502
 503
 504
 505
 506
 507
 508
 509
 510
 511
 512
 513
 514
 515
 516
 517
 518
 519
 520
 521
 522
 523
 524
 525
 526
 527
 528
 529
 530
 531
 532
 533
 534
 535
 536
 537
 538
 539
 540
 541
 542
 543
 544
 545
 546

DU 221 GGIGAAATATGCACATCAAAGAAGAGAAATTTTTCACCTTGACAGCCCTATGCTGTTCCATG 280

[illegible]

QY 261 taccctggaggatctcgaccgggagttcaaggccttttggcccatgctcaat---ggaaagc 317

281 282 283 284 285 286 287 288 289 290 291 292 293 294 295 296 297 298 299 300 301 302 303 304 305 306 307 308 309 310 311 312 313 314 315 316 317 318 319 320 321 322 323 324 325 326 327 328 329 330 331 332 333 334 335 336 337 338 339 340 341 342 343 344 345 346 347 348 349 350 351 352 353 354 355 356 357 358 359 360 361 362 363 364 365 366 367 368 369 370 371 372 373 374 375 376 377 378 379 380 381 382 383 384 385 386 387 388 389 390 391 392 393 394 395 396 397 398 399 400 401 402 403 404 405 406 407 408 409 410 411 412 413 414 415 416 417 418 419 420 421 422 423 424 425 426 427 428 429 430 431 432 433 434 435 436 437 438 439 440 441 442 443 444 445 446 447 448 449 450 451 452 453 454 455 456 457 458 459 460 461 462 463 464 465 466 467 468 469 470 471 472 473 474 475 476 477 478 479 480 481 482 483 484 485 486 487 488 489 490 491 492 493 494 495 496 497 498 499 500 501 502 503 504 505 506 507 508 509 510 511 512 513 514 515 516 517 518 519 520 521 522 523 524 525 526 527 528 529 530 531 532 533 534 535 536 537 538 539 540 541 542 543 544 545 546 547 548 549 550 551 552 553 554 555 556 557 558 559 560 561 562 563 564 565 566 567 568 569 570 571 572 573 574 575 576 577 578 579 580 581 582 583 584 585 586 587 588 589 590 591 592 593 594 595 596 597 598 599 600 601 602 603 604 605 606 607 608 609 610 611 612 613 614 615 616 617 618 619 620 621 622 623 624 625 626 627 628 629 630 631 632 633 634 635 636 637 638 639 640 641 642 643 644 645 646 647 648 649 650 651 652 653 654 655 656 657 658 659 660 661 662 663 664 665 666 667 668 669 670 671 672 673 674 675 676 677 678 679 680 681 682 683 684 685 686 687 688 689 690 691 692 693 694 695 696 697 698 699 700 701 702 703 704 705 706 707 708 709 710 711 712 713 714 715 716 717 718 719 720 721 722 723 724 725 726 727 728 729 730 731 732 733 734 735 736 737 738 739 740 741 742 743 744 745 746 747 748 749 750 751 752 753 754 755 756 757 758 759 760 761 762 763 764 765 766 767 768 769 770 771 772 773 774 775 776 777 778 779 780 781 782 783 784 785 786 787 788 789 790 791 792 793 794 795 796 797 798 799 800 801 802 803 804 805 806 807 808 809 810 811 812 813 814 815 816 817 818 819 820 821 822 823 824 825 826 827 828 829 830 831 832 833 834 835 836 837 838 839 840 841 842 843 844 845 846 847 848 849 850 851 852 853 854 855 856 857 858 859 860 861 862 863 864 865 866 867 868 869 870 871 872 873 874 875 876 877 878 879 880 881 882 883 884 885 886 887 888 889 890 891 892 893 894 895 896 897 898 899 900 901 902 903 904 905 906 907 908 909 910 911 912 913 914 915 916 917 918 919 920 921 922 923 924 925 926 927 928 929 930 931 932 933 934 935 936 937 938 939 940 941 942 943 944 945 946 947 948 949 950 951 952 953 954 955 956 957 958 959 960 961 962 963 964 965 966 967 968 969 970 971 972 973 974 975 976 977 978 979 980 981 982 983 984 985 986 987 988 989 990 991 992 993 994 995 996 997 998 999 1000 1001 1002 1003 1004 1005 1006 1007 1008 1009 1010 1011 1012 1013 1014 1015 1016 1017 1018 1019 1020 1021 1022 1023 1024 1025 1026 1027 1028 1029 1030 1031 1032 1033 1034 1035 1036 1037 1038 1039 1040 1041 1042 1043 1044 1045 1046 1047 1048 1049 1050 1051 1052 1053 1054 1055 1056 1057 1058 1059 1060 1061 1062 1063 1064 1065 1066 1067 1068 1069 1070 1071 1072 1073 1074 1075 1076 1077 1078 1079 1080 1081 1082 1083 1084 1085 1086 1087 1088 1089 1090 1091 1092 1093 1094 1095 1096 1097 1098 1099 1100 1101 1102 1103 1104 1105 1106 1107 1108 1109 1110 1111 1112 1113 1114 1115 1116 1117 1118 1119 1120 1121 1122 1123 1124 1125 1126 1127 1128 1129 1130 1131 1132 1133 1134 1135 1136 1137 1138 1139 1140 1141 1142 1143 1144 1145 1146 1147 1148 1149 1150 1151 1152 1153 1154 1155 1156 1157 1158 1159 1160 1161 1162 1163 1164 1165 1166 1167 1168 1169 1170 1171 1172 1173 1174 1175 1176 1177 1178 1179 1180 1181 1182 1183 1184 1185 1186 1187 1188 1189 1190 1191 1192 1193 1194 1195 1196 1197 1198 1199 1200 1201 1202 1203 1204 1205 1206 1207 1208 1209 1210 1211 1212 1213 1214 1215 1216 1217 1218 1219 1220 1221 1222 1223 1224 1225 1226 1227 1228 1229 1230 1231 1232 1233 1234 1235 1236 1237 1238 1239 1240 1241 1242 124

DB 281 GACCCAGAGGAAATTGTGATCGCGTACGCTGGGCTACACTCAAGGGTCTCTTTGAAACAGA 340

[illegible]

318 acaa^gac^gaa^gta^tat^tct^cta^tat^tg^ggt^cata^caa^tg^aca^ttt^tg^act^tat^t 377

341 342 343 344 345 346 347 348 349 350 351 352 353 354 355 356 357 358 359 360 361 362 363 364 365 366 367 368 369 370 371 372 373 374 375 376 377 378 379 380 381 382 383 384 385 386 387 388 389 390 391 392 393 394 395 396 397 398 399 400 401 402 403 404 405 406 407 408 409 410 411 412 413 414 415 416 417 418 419 420 421 422 423 424 425 426 427 428 429 430 431 432 433 434 435 436 437 438 439 440 441 442 443 444 445 446 447 448 449 450 451 452 453 454 455 456 457 458 459 460 461 462 463 464 465 466 467 468 469 470 471 472 473 474 475 476 477 478 479 480 481 482 483 484 485 486 487 488 489 490 491 492 493 494 495 496 497 498 499 500 501 502 503 504 505 506 507 508 509 510 511 512 513 514 515 516 517 518 519 520 521 522 523 524 525 526 527 528 529 530 531 532 533 534 535 536 537 538 539 540 541 542 543 544 545 546 547 548 549 550 551 552 553 554 555 556 557 558 559 560 561 562 563 564 565 566 567 568 569 570 571 572 573 574 575 576 577 578 579 580 581 582 583 584 585 586 587 588 589 590 591 592 593 594 595 596 597 598 599 600 601 602 603 604 605 606 607 608 609 610 611 612 613 614 615 616 617 618 619 620 621 622 623 624 625 626 627 628 629 630 631 632 633 634 635 636 637 638 639 640 641 642 643 644 645 646 647 648 649 650 651 652 653 654 655 656 657 658 659 660 661 662 663 664 665 666 667 668 669 670 671 672 673 674 675 676 677 678 679 680 681 682 683 684 685 686 687 688 689 690 691 692 693 694 695 696 697 698 699 700 701 702 703 704 705 706 707 708 709 710 711 712 713 714 715 716 717 718 719 720 721 722 723 724 725 726 727 728 729 730 731 732 733 734 735 736 737 738 739 740 741 742 743 744 745 746 747 748 749 750 751 752 753 754 755 756 757 758 759 760 761 762 763 764 765 766 767 768 769 770 771 772 773 774 775 776 777 778 779 780 781 782 783 784 785 786 787 788 789 790 791 792 793 794 795 796 797 798 799 800 801 802 803 804 805 806 807 808 809 810 811 812 813 814 815 816 817 818 819 820 821 822 823 824 825 826 827 828 829 830 831 832 833 834 835 836 837 838 839 840 841 842 843 844 845 846 847 848 849 850 851 852 853 854 855 856 857 858 859 860 861 862 863 864 865 866 867 868 869 870 871 872 873 874 875 876 877 878 879 880 881 882 883 884 885 886 887 888 889 890 891 892 893 894 895 896 897 898 899 900 901 902 903 904 905 906 907 908 909 910 911 912 913 914 915 916 917 918 919 920 921 922 923 924 925 926 927 928 929 930 931 932 933 934 935 936 937 938 939 940 941 942 943 944 945 946 947 948 949 950 951 952 953 954 955 956 957 958 959 960 961 962 963 964 965 966 967 968 969 970 971 972 973 974 975 976 977 978 979 980 981 982 983 984 985 986 987 988 989 990 991 992 993 994 995 996 997 998 999 1000 1001 1002 1003 1004 1005 1006 1007 1008 1009 1010 1011 1012 1013 1014 1015 1016 1017 1018 1019 1020 1021 1022 1023 1024 1025 1026 1027 1028 1029 1030 1031 1032 1033 1034 1035 1036 1037 1038 1039 1040 1041 1042 1043 1044 1045 1046 1047 1048 1049 1050 1051 1052 1053 1054 1055 1056 1057 1058 1059 1060 1061 1062 1063 1064 1065 1066 1067 1068 1069 1070 1071 1072 1073 1074 1075 1076 1077 1078 1079 1080 1081 1082 1083 1084 1085 1086 1087 1088 1089 1090 1091 1092 1093 1094 1095 1096 1097 1098 1099 1100 1101 1102 1103 1104 1105 1106 1107 1108 1109 1110 1111 1112 1113 1114 1115 1116 1117 1118 1119 1120 1121 1122 1123 1124 1125 1126 1127 1128 1129 1130 1131 1132 1133 1134 1135 1136 1137 1138 1139 1140 1141 1142 1143 1144 1145 1146 1147 1148 1149 1150 1151 1152 1153 1154 1155 1156 1157 1158 1159 1160 1161 1162 1163 1164 1165 1166 1167 1168 1169 1170 1171 1172 1173 1174 1175 1176 1177 1178 1179 1180 1181 1182 1183 1184 1185 1186 1187 1188 1189 1190 1191 1192 1193 1194 1195 1196 1197 1198 1199 1200 1201 1202 1203 1204 1205 1206 1207 1208 1209 1210 1211 1212 1213 1214 1215 1216 1217 1218 1219 1220 1221 1222 1223 1224 1225 1226 1227 1228 1229 1230 1231 1232 1233 1234 1235 1236 1237 1238 1239 1240 1241 1242 1243 1244 1245 1246 1247 1248 1249 1250 1251 1252 1253 1254 1255 1256 1257 1258 1259 1260 1261 1262 1263 1264 1265 1266 1267 1268 1269 1270 1271 1272 1273 1274 1275 1276 1277 1278 1279 1280 1281 1282 1283 1284 1285 1286 1287 1288 1289 1290 129

DB 341 ACATCTCTGAAGAGATATCTAGAAAGTATGGCAATATATGAACAAATGTAATCTTTGGCCCT 400

Qv 378 + + + + + c a a a a + + + c a a a a a + + + + + 377

WY 378 CCCCCCAATCCAGGAGCCCTCCTAAGACCAAAAATCTGTAGAATACTTAAAGGAGAG 437

db 401 TCATAGGTCTGTGGAGCTATCTGGCCCTCATTCAGCCACCTCATTCCTAC 160

DD 1CATATGGTGTGGTGGAGCTACCTGCAATAAAGHGGCCCTGGATCAGGCACCTGGAATGCTAC 460
 401

[illegible]

QY 846 ctgcatcagggaacgctgctgctggaattgaagcctacattaatgcttctgaga 905
 Db 481 CTCTCATCAGGGAAGCAATGCTATGGAATTTGAAGCCTACATTAATGCTCTGAGA 540
 QY 906 acatggaattggtttctcttggatcaaatgctcagaataatcagaagaagaagc 965
 Db 541 ACA:GGAATGTGGTTTCTTTGGATCAATGGTCTCAGAAATTCAGAGAAGAAGC 600
 QY 966 tatggaattgctgctgcttgggcaaatccctcagacagctgctgctgctgctgctg 1025
 Db 601 TATGGAATTTGCTGATGCTTTGGGCAAAATCCCTCAGACAGTCTCTGGCGGTACACTGG 660
 QY 1026 aaccgaccatcgaattctgcaacaacacgatactgttaagtgcctaccaccaaaacga 1085
 Db 661 AACCCGACCATCGAATCTTCGGAACAACACGATCTTTGAAGTGGCTACCCCAACGA 720
 QY 1086 tctgctgtgtaaccgctgaccgctgctcttatcaccatgctggttcccatggtttta 1145
 Db 721 TCTGCTGTGTACCCGATGACCCGCTGCTTTATCACCATGCTGCTGCTGCTGCTGCTTA 780
 QY 1146 tgaagcatatgcaatggcttcccatgctgctgctgctgctgctgctgctgctgctg 1205
 Db 781 TGAAGCATATGCAATGGCTTCCCATGCTGATGATGCTGCTGCTGCTGCTGCTGCTGA 840
 QY 1206 caatgcaagcgcatgagactaaggagctggagtgacccctgaatgcttctggaatgac 1265
 Db 841 CAATGCAAGCGCATGAGACTAAGGAGCTGGAGTCAACCTGAATGCTTGGAAATGAC 900
 QY 1266 tctggaattgaaatgctctaaagcagctcatcaatgacaaaagtacaaaggagaa 1325
 Db 901 TTTCTGAAGATTAGAAAATGCTCTAAAGCAGTCTCATCAATGACAAAAGTTACAGAGGAA 960
 QY 1326 catcatgagccttccagccttcaaaagcagcgcgcgcgcgcgcgcgcgcgcgcgcgc 1385
 Db 961 CATCATGCGCTCTCCAGCCTTCAAGAGCCGCGCGGTGAGCCGCTGGACCTGGCCGT 1020
 QY 1386 gtctggtggtggttggatgagcagcagcagcagcagcagcagcagcagcagcagcagc 1445
 Db 1021 GTTCTGGGTGGAGTTTGTGATAGGCACAAAGGCGCGCCACACCTGCGCCCGCAGCCCA 1080
 QY 1446 cgactccactcgttaccagctacattctctgagctgagctgagctgagctgagctgagct 1505
 Db 1081 CGACCTACCTGTTACCACTGACCTGCTTGGAGTGTGCTTGGCGCGCTGCT 1140
 QY 1506 gctgacagtgcccttcacattcaattggttggcttaccgagcagcagcagcagcagcagc 1565
 Db 1141 GCTGACGTGCGCTTCACTACCTTTAAATGTTGTGCTTATGCTACCGGAAATGCTTGGG 1200
 QY 1566 gaaaaagggcgagtttaagaagcccaacaatccagagccattgagaagtgagtgagaa 1625
 Db 1201 GAAAAAGGCGAGTTAAGAAAGCCCAAAATCCAGAGCCCATTTGAGAGTGGGTGGAA 1260
 QY 1626 taaggttaaaatttgaacattccctgctgcttcccaacttgaacagagaatcagtgct 1685
 Db 1261 ATAAGTAAAAATTTGAACCATTCCTGATCAATTTCAAACTTGAACACAGAAATCAAGTGT 1320
 QY 1686 taattcatttttcttcttaaggaatacttgcataaataatcagcagcagcagcagcagc 1745
 Db 1321 TAAATTCATTTTATCTTATTAGGAATACCTTGGCTAAATTAATACGCCCCAGAGTGC 1380
 QY 1746 ttttaaaattctcttaataaaataatagactcgtgctgctgctgctgctgctgctgctgct 1805
 Db 1381 TTTAAAAAATTTCTTTAAAAATAAATAATAGACTCGCTAGTCAAGTAAAGATATTGGAATA 1440
 QY 1806 tgtatcgtgccccctcgcgttcttgcatacagatgacatgctgcttcttctcagagacg 1865
 Db 1441 TGTATCGTCCCCCTCTGCTGCTTTTGTATCAGATGACATGCTGCTTTTTCAGAGGACG 1500
 QY 1866 tgcagacagcgtgagcttctgattacttcttcttcttcttcttcttcttcttcttcttct 1925
 Db 1501 TGCACACAGGCTGCATCTTAGATTACTTTTCTTACTTGAACATGCGCTGTTTGGAG 1560
 QY 1926 tgcgggattcaagggttgcctccacgcctgcccctactgcaaatggcagtttcttcttct 1985

Db 1561 TGGGGATTCAAAGGTGGTCCACGCTGCCCTACTGCAATGGCAGTTTAATCTTAT 1620
 QY 1986 ctttggctctgcagatggttgcataatgcttcaaccataatggtcagtcctcatct 2045
 Db 1621 CTTTTGGCTTCGAGATGGTTGCAATTCATCTTAACCAATAATGGTCAGTCTCATCT 1680
 QY 2046 ctgctcctcttcataaggctgacccctgctgctttaaagaaggaagccttcttcttcttag 2105
 Db 1681 CTGCTGCTGTTTCATAGGTGCCCTTGTGTTTAAAGAAGGAGCTTTGACCTTTAG 1740
 QY 2106 agtgtagtgtaaatgtaagtgaatgcttgcctgagcagcagcagcagcagcagcagcagc 2165
 Db 1741 AGTGTAGGTGAATGAATGAATGGCTTGGAGTGCACCTGAGAACAGCATATGATTTCTTGC 1800
 QY 2166 ttgggggaaaaaagaatgatgctatgaaatggtggtggtggtggtggtggtggtggtggtggt 2225
 Db 1801 TTTGGGGA-AAAAGAATGATGCTATGAAATTTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1860
 QY 2226 tgcctatg-caaatggaactgaattgataaaaaaaccccaaacacacagcctatgaaagctgctg 2285
 Db 1861 TGCCTATGTCAAATGGAGCTGAAATTTGATAAAACCCCAAAATACAGCTATGAAGTCTGG 1920
 QY 2286 gcaagttacttttttctgctgcttctcctcaact 2320
 Db 1921 GCAAGTTTACTTTTTTCTGATGTTTCTCTACAAC 1955

RESULT 10
 HUMHUGBR2 2368 bp mRNA linear PRI 07-MAR-1995
 LOCUS Human bilirubin UDP-glucuronosyltransferase isozyne 2 mRNA, complete cds.
 DEFINITION M57951.1 GI:184474
 ACCESSION M57951
 VERSION 1
 KEYWORDS UDP-glucuronosyltransferase; bilirubin UDP-glucuronosyltransferase.
 SOURCE Human adult female liver, CDNA to mRNA, clones Z17 and Z17P5E2.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 RITTER, J.K., Crawford, J.M. and Owens, I.S.
 Cloning of two human liver bilirubin UDP-glucuronosyltransferase cDNAs with expression in COS-1 cells
 J. Biol. Chem. 266 (2), 1043-1047 (1991)
 JOURNAL 91093210
 MEDLINE Location/Qualifiers
 FEATURES
 source
 1. .2368
 /organism="Homo sapiens"
 /isolate="AK"
 /db_xref="taxon:9606"
 /map="Unassigned"
 /clone="Z17"
 /sex="female"
 /tissue_type="liver"
 /dev_stage="adult"
 /tissue_lib="lambda-ZAP"
 1. .2368
 /gene="UGT2B"
 30. .1634
 /gene="UGT2B"
 /standard_name="bilirubin UDP-glucuronosyltransferase
 isozyne 2"
 /BC_number="2.4.1.17"
 /codon_start=1
 /product="UDP-glucuronosyltransferase 2 family polypeptide B"
 /protein_id="AAA63196.1"
 /db_xref="GI:184475"
 /db_xref="GDB:G00-127-753"
 /translation="MARGLOVPLPRLATGLILLISVOPWAESGKVLVVPDGSPLWSM
 REALRELHARGHOAVLITPEVNHKIEKFFLTATAYPTQKEFDRTVLTGTTQGEFFE
 THELLKRYSRSMIMNNVSLAHLRCCVLELLHNEALIRHLNATYSDVLTDPVLCGAV

gene

CDS


```

VERSION AF104336.1 GI:6537137
KEYWORDS crab-eating macaque.
SOURCE Macaca fascicularis
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Cercopitheciidae;
Cercopitheciinae; Macaca.
REFERENCE 1 (bases 1 to 2396)
AUTHORS Albert, C., Vallee, M., Beaudry, G., Belanger, A. and Hum, D.W.
TITLE The monkey and human uridine diphosphate-glucuronosyltransferase
UGT1A9, expressed in steroid target tissues, are
estrogen-conjugating enzymes
JOURNAL Endocrinology 140 (7), 3292-3302 (1999)
MEDLINE 99312152
PUBMED 10385426
REFERENCE 2 (bases 1 to 2396)
AUTHORS Albert, C., Vallee, M., Beaudry, G., Belanger, A. and Hum, D.W.
TITLE Direct Submission
JOURNAL Submitted (05-NOV-1998) Molecular Endocrinology, Laval University,
2705 Laurier Boulevard, Sainte-Foy, Que G1V4G2, Canada
FEATURES
source
Location/Qualifiers
1..2396
/organism="Macaca fascicularis"
/db_xref="taxon:9541"
60..1652
/codon_start=1
/product="UDP-glucuronosyltransferase UGT1A9"
/protein_id="AAFI5546.1"
/db_xref="gi:6537138"
/translation="MTRTGWSPLPLCVCLLLAGFAEAGKLVVVPDGSWLSMRSV
VEKLIRGHEVVVMPKVSQWOLRPLNCTVKTYSYTLIEDLDRFAFRQAOAKQV
QSIFSLIMSSYNDIFLFFSNCRSLFKDKLVLEKSFSDAVFLDPDTCGLIVAKY
LSLPSWFAFGVCHYLEGACOPAPLSYPRILLGFSDAWTFKRVNHLMLERL
LCRFEKSALEISEILOTPVVDLHROLSIMLLRNDVFLDYFKVPWPNWTFGIN
CHQGPLPMEFEAYINASGEHIVFSLGSMVAIEFKKMAIALADALQIPQVLWRY
TCFPPSNLANTILVWLQPNDLIGHMPNRAFITHAGSHGIYEGICNGVPMWPMPLFG
DOMDNKRMETKAGVTLNVLEMTSEDLNALKAVINDKSYKENIMHLSLHRDREVE
PDLAVFVWFVPMHKGAPHLRPAHDLTWYQVHSLDVIFGLLAILVTVAIFAKCCA
YGYRCKFGKGRVKRAHKSKTH"
BASE COUNT 645 a 513 c 558 g 680 t
ORIGIN

Query Match 89.5%; Score 2076; DB 9; Length 2396;
Best Local Similarity 94.6%; Pred. No. 0;
Matches 2194; Conservative 0; Mismatches 120; Indels 6; Gaps 4;

QY 2 teagtcaggtctctgtatggtgtgacaggggtgacagccccccttctctatgtgt 61
DB 43 TCGGCTCGGTTCCCTGTATGACTCGCACAGGGTGGAGCAGTCCCTTCCCTTATGTGT 102

QY 62 gtctgtctgacgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 121
DB 103 GTCTGCTGCTGGCTGTGGCTTTGCCAGGCGAGGAGCTGCTGGTAGTGGCCGCGGATG 162

QY 122 ggagccactgttccactatgaggtcggtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 181
DB 163 GGAGCCACTGGCTCTCCATAGAGTCGGTGTGGTGGAGAACTCATCTCTCAGGGGCGATGAG 222

QY 182 Tggtgtgtagtcagagaggtgagttggaactgttccaggggaggttcaaggctttggcc 241
DB 223 TGGTGTGAGTCATGCAAGGTGAGTTGGCAGCTGGGAGAGCACTCATCTCTCAGGGGCGATGAG 282

QY 242 agactatcaactcaatataccctggaggtatctgacggggaggttcaaggctttggcc 301
DB 283 AGACTTATTCACACTTATACACTCTGGAGGATCTGGACGGGAGTTCAAGGCTTTTCGCC 342

QY 302 atgctcaatgaaagcaagtcagagtatatactctctatgaatggttcatacaatg 361
DB 343 AGGCTCAATGGAAGCAGACAGTACAAAGTATATTTCTTAATANTAGTTCCTACATG 402

QY 362 acattttgactatttttcaaatgcaggaggtgtgtgtgtgtgtgtgtgtgtgtgtgt 421
DB 1483 CCCAGGACCTCACCTGGTACCAGTACCATTCTCTGGAGGTGATCGGTTTCCTCTCGCCA 1542

```


Qy	2284	ggcgaagttaacttttttctatgattcttcacaaact	2320
Db	2336	GGCCAAGTTTACTTTTTTCTGATGTTCCTACAACT	2372
RESULT	7		
S55985			
LOCUS		S55985	2392 bp mRNA linear PRI 04-MAY-2000
DEFINITION		S55985	Homo sapiens UDP-glucuronosyltransferase mRNA, complete cds.
ACCESSION		S55985	
VERSION		S55985.1	GI:235453
KEYWORDS			
SOURCE			human liver.
ORGANISM			Homo sapiens
REFERENCE			Mammalia; Eutheria; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. 1 (bases 1 to 2392) Wooster,R., Sutherland,L., Ebner,T., Clarke,D., Da Cruz e Silva,O.A. and Burchell,B. Cloning and stable expression of a new member of the human liver phenol/bilirubin: UDP-glucuronosyltransferase cDNA family Biochem. J. 278 (pt 2), 465-469 (1991)
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
REMARK			GenBank staff at the National Library of Medicine created this entry [NCBI gbbseq 55985] from the original journal article.
FEATURES			
source			1..2392
		/organism="Homo sapiens"	
		/db_xref="taxon:9606"	
		/tissue_type="liver"	
CDS		35..1627	
		/note="phenol/bilirubin; lugP4; conceptual translation presented here differs from translation in publication" /codon_start=1 /product="UDP-glucuronosyltransferase" /protein_id="AAB19791.2" /db_xref="GI:7690346" /translation="MACTGWTSPLCPVCLLLTCTCGFAEAGKLVVVPMDGSWFTMRSV VEKLIIRGHEVYNDIAPLFSNCRSLFKOKKLVEYLKESSFDADFDPDNCGLLIYAKY RSIYLLSGMSVDIPDLFSGNARSLEPKKKLYELVESFDVAFDFPDNCGLLIYAKY FSLPSVFRGILCHILEGAQPAPLSYVRPLLGSDDMTFKERVNRNHIMHLEEHL LCRRFLPMEEAIENLIQTPTVEDLYSHTSIWLLRTDFVLDYPRKVPMPNNIFIGGIN CHERALKMEFEAYINASGEHGIVGSIKSWSEIKKAMAIALDAKIPTVLMRY TRTCRNLANETILVLKPNOLLGHLPMTSAFIETHAGSHGVYESICNGPVMMWMLPG DQMDNAKMRETKGATGYTMNYLENTSELENALKAVINDKSYKENIMRLSSLHKDRPVE PIDLAFWVEVMHKGATHLRPAADHLTWYQHSLDVIGFLAAVLTVAFITCKCCA YGRXCLCKGRVKKHAHSKTH"	
BASE COUNT		661 a 492 c 554 g 685 t	
ORIGIN			
Query Match		93.0%; Score 2156.8;	DB 9; Length 2392;
Best Local Similarity		97.7%;	Pred. No. 0;
Matches 2362;		Conservative 0;	Mismatches 47; Indels 7; Gaps 7;
Qy	9	cagttctctaagctgcacagggtggaccagccccccttccttatgtgtgtcgtgt	68
Db	25	CACCTCTGGCATGCCTTGCCAGGGTGACACGCCCTTCCTCTATGTGTGTCTGCT	84
Qy	69	gctgaccttggtcttcgcaggcagggaagactactgtagtgcacctgatggaggcca	128
Db	85	GCTGACCTTGCGCTTTCGCGAGCGAGGAAAGCTAGTGTGTGTGCCATGGAGCCA	144
Qy	129	ctgggttcaccatgagctcggttgttgagaacatcatctcagggggcatagatgattg	188
Db	145	CTGGGTTCACCATGAGGTGCGGTGGTGGAGAACAATCATCTCAGGGGCATGAGGTGTGT	204
Qy	189	agctca-tgccagggtgadtgtggacaacttggaagatcactgaattgcacagtgaactt	247
Db	205	AGTCAGTGCCACAGAGGTGAGTGTGGCACTGGGSAAGATCAATGAATTGACACAGTGAACACTT	264
Qy	248	attcaacttcatacaccttqqaqatctcqaccqqgaqtccaqccttttgcaccaatgctc	307

DEFINITION	<p> Homo sapiens, Similar to UDP glycosyltransferase 1 family, polypeptide A9, clone MGC:9353 IMAGE:3850654, mRNA, complete cds. BC020971 BC020971.1 GI:18088086 MGC. SOURCE human. ORGANISM Homo sapiens Eukaryota; Metazoa; Chordata; Cranita; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo. Strausberg, R. Direct Submission Submitted (03-JAN-2002) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA </p>
REMARK	<p> NIH-MGC Project URL: http://mgc.nci.nih.gov Contact: MGC help desk Email: cgapbs-r@mail.nih.gov Tissue Procurement: ATCC CDNA Library Preparation: Life Technologies, Inc. CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLML) DNA Sequencing by: Sequencing Group at the Stanford Human Genome Center, Stanford University School of Medicine, Stanford, CA 94305 Web site: http://www-shgc.stanford.edu Contact: (Dickson, Mark) mcdepaxil@stanford.edu Dickinson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers, R. M. </p>
COMMENT	

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
 Series: TRAK Matrix: 20 Row: 8 Column: 8
 This clone was selected for full length sequencing because it passed the following selection criteria: Genomescan gene prediction, Similarity but not identity to protein.

```

FEATURES
Location/Qualifiers
1. .2423
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="MGC:9353 IMAGE:3850654"
/class="type=Colon, adenocarcinoma"
/classifier="NH, MGC_65"
/lab_host="DH10B"
/notes="Vector: pCMV-Sport6"

```

```

70..1662      /codon_start=1
                /product="similar to UDP glycosyltransferase 1 family,
polypeptide A9"
                /protein_id="AAH20971.1"
                /db_xref="GI:18080807"
                /translation="MARAGQWSPVLCVLLTCGFAEAGKLLVPMQGSWFFTMQSV
VEKLLIRGHEVVVYVMEVSWQLERSLNCTVKYYSSTYLEDQQRNEFMFAHQWKAQA
QSIFSLWSSSGSLFDHFSRCLFNDNRKLVEYLKESSEYTLDPDFTDGLGLVARY
FSLSPVTRIFGCHLEAGCAQAPLSPVNDLLGFSADAMTEKRVNWNHVLHLEDL
FCQYIFRNALTEISLQTPVTDLYLSHTLSWLLRTDVLIDYPKPMNNIFITGGIN
CHQGPLMEAEAYINASEGIEIVFVSLGMSYEIPEKKAMATADALGKIQTIVLYIN
TGTSPNSLANNTLLVKWLPQNDLLGHWPTEARTHAGSHGVYESICNGVPMVWMPLEF
QDMQNDKARMEYKAGVTLNLMESIDENALKAINDSKENIMRUSLKKHDKORPVE
PDLAVFWVEVMHKGAPLHRAAHDLTWYQTHSLDVIGFLAVLVITVAITFKGCA
YGRKCLAKGRVKKAHKSHTH"
BASE COUNT      643 a      514 c      707 t

```

```

Query Match      94.6%; Score 2194.6; DB 9; Length 2423;
Best Local Similarity 97.0%;
Matches 2247; Conservative 0; Mismatches 69; Indels 1; Gaps 1;

OY 5 gctgcagttctctgatggtgtgcacaggggtggaccagcccctctctatgtgtgtc 64
Db 56 GCTGCAGTTCTCTCATGCTCCGCGAGGGTGGACCGCCGCTTCTTATGTGTGTC 115
OY 65 tgcTgcTgcacctgtggttttgcaggcaggggaagctactgttagtgcctatgatggga 124

```

Db	116	TACTGCTGACCTGTGGCTTTGGCGAGGACGAGGAAGTGTGCTAGTCCCATGGATGGGA	175
QY	125	gccactgggttcaccatgagtgcggtgggtgggaaactcattctcagggggcagtaggggg	184
Db	176	GTCACTGGTTACACATGCAGTCGGTGGTGAGAAACTTATCTCAGGGGGCATGAGGTGG	235
QY	185	tgttagtcatgcacagaggtgagttggcaactgggaagatcactgaattgcacagtgaaga	244
Db	236	TTGTAGTTCATGCCAGAGGTGAGTGGCAACTGGAAAGATCACTGAATGACACAGTAA	295
QY	245	cttattcaactcatataccctggagagatctggacgggagttcaaggcttttgcaccatg	304
Db	296	CTTACTCAACCTCGTACACTCTGGAAGATCAGAACCGGGAATTCATGGTTTGCGCCATG	355
QY	305	ctcaatggaaagcacacagtcagagtatatatctctatataatgggtttcatacaatgaca	364
Db	356	CTCAATGGAAAGCACAGGCACAAGTATATTTCCTATTAAATGAGTTCATCCAGTGGTT	415
QY	365	tttttgactatttttttcaaatgcaggagttgttttaagacacaaaaattagtagaat	424
Db	416	TTCTTGACTATTATTTTCCGATTCGCGAGTGTGTGTTAATGACCGAANAATAGTAGAAT	475
QY	425	acttaagagagagttcttttgatgagtgagttctctcatctctttgatcaacttgcgttaa	484
Db	476	ACTTAAAGAGAGAGTTCCTTTGATGTCAGTGTTCGTGATGCTTTTGATACCTGTGCTTAA	535
QY	485	ttgtggcaaaattttctccctcccctcgttggtcttcgcaggggaatactttgcacct	544
Db	536	TTGTGTCAAAATATTTCTCCTCCCTCTGTGGTCTCACCAGGGGAATATTTTGCACCC	595
QY	545	atcttgaaagagtgacagtgccctgctctctcttctctatgtccccagaattctcttag	604
Db	596	ATCTTGAAGAAGGTGCACAGTGCCTCTCTCTTCTTCCTATGTCGCCAATGATCTCTTAG	655
QY	605	ggttctcagatgccatgactttcgaaggagagtagcgaaccacatcatgctacattggagg	664
Db	656	GGTTCTCAGATGCCATGACTTTTCAAGGAGAGAGTATGGAACACACATGCTGCACTTGGAGG	715
QY	665	aacatttatggcaacggttttttcaaaatgccttagaataagcctctgaaattctcc	724
Db	716	ACCATTATTTTGGCAGTATCTTTTGTAGAAATGCCCTAGAAATAGCCTCTGAAATCTCC	775
QY	725	aaacacgtgtacggagtagtatctctacagcaacacatacaatttggtgtgtcgaaacgg	784
Db	776	AAACCCCTGTACGCGATATGATCTCTACATGCACATCAATTTGGTTGTGCGAAACGG	835
QY	785	acttggtttggactatcccaaacccgtgatgcccaacatgatctcattggttggtatca	844
Db	836	ACTTTGTTTGGACTATCCCAACCCGTGATGCCAACATGATCTTCATTGGTGGTATCA	895
QY	845	actgcaatcagggaagccgttgctatggaattgaaacctataatgaattgtcttgag	904
Db	896	ACTGTATCAGGAAAGCCATAGCCTATGGAATTTGAAAGCCTACATTAATGCTCTTGAG	955
QY	905	aacatgaaattgggttttctctttgggatcaaaggcttcagaaattccagagaagaag	964
Db	956	AACATGGAATTTGTGTTTCTCTTTTGGGATCAATGTGCTCAGAAATCCAGAGAAGAA	1015
QY	965	ctatggaattgtcgtatgctttgggcaaaatccctcagacagtcctgtggcggtacactg	1024
Db	1016	CTATGGCAATTCGTATGCTTTTGGGCAAAATCCCTCAGACAGTCTGTGGCGGTACACTG	1075
QY	1025	gaaccggacactgaaattcttgcgaacacacgatacttgttaagtggctaccaccaaacg	1084
Db	1076	GAACCCGACCATGAAATCTTTCGGAACACACAGATPACTTGTTAAGTGGCTACCCCAAAACG	1135
QY	1085	atctgcttggtcacccatgacccgtgctttatcacccatgctgggttcccactggtttt	1144
Db	1136	ATCTGCTTGGTCACCCGATGACCCGTGCTTTATACCCATGCTGGTGTCCCATGTGTTTT	1195
QY	1145	atgaagcatatgcaatggcggttcccatcgtgtgatgccttgttgggtatcagatgg	1204

Qy 1205 acaatgcgaagcgcgtgagactaagagctggagctgagctgacccctgaatgtcttctggaatga 1264
|||||
Db 1220 ACAATGCAAAAGCGATGGAGCTAAGGGAGCTGGAGTGCACCTGAATGCTTCTGGAATGA 1279
Qy 1265 ctctgaagattagaataatgctctaaagcagtcacatcaatgacaaagttacaagaga 1324
Db 1280 CTTCTGAGATTTAGAAATGCTCTAAAGCAGTCAATCAATGACAAAGTTACAAGAGA 1339
Qy 1325 acatcatgcgctctccagcctcaagcctcaagagccgcgctgagcgcgtggaacctggccg 1384
Db 1340 ACATCATGCGCTCTCCAGCCTTCACAAGGACCGCCGCTGGAGCGCTGGAGCTGGCGG 1399
Qy 1385 tgtctgggtgggttggatgagggcacaaggcgccacacactgcgcccgcagccc 1444
Db 1400 TGTCTGGGTGGAGTTGTGTGATGAGGACAAAGGCGGCCACACCTGGCCCGCCGAGCCC 1459
Qy 1445 acgacctcaactgtgaccagtaaccattccttggacgtgattgttctctcttggccgtcg 1504
Db 1460 ACGACCTCACCTGCTACAGTACCATTCCTTTGGACGTGATTGTTCTCTTGGCGGTGG 1519
Qy 1505 tgcgcagctggccttcacacattcaaatgtgtcttatgctaccggaatgcttg 1564
Db 1520 TGTGACAGTGGCCTTCATCACCCTTAATGCTGTCTTATGGCTACCGGAAATGCTTG 1579
Qy 1565 ggaataagggcggttaagaagcccaataatccaaagcccaattggaagtggtggga 1624
Db 1580 GGAATAAGGGCGAGTTAAGAAGGCCCAAAATCCAAAGACCCATTTGAGAAGTGGGTGG 1639
Qy 1625 aataaggtataaatttgaacattccctagtcatttccaaacttgaataacagacatcagtg 1684
Db 1640 AATAAGGTAAATTTGAACCATTCCTTAGTCATTTCCAACTTCAAAACAGAAATCAGTG 1699
Qy 1685 ttaattcattttatttatttaaggaataacttgcataaataatcaagccccagtg 1744
Db 1700 TTAATTTCAATTTATTTAATAAGGAATAACTTTGCATAAAATTAATCAGCCCCAGAGTG 1759
Qy 1745 ctttaaaaaattctttaaataaaaaataagactgcagtcagtaaaagatttgaat 1804
Db 1760 CTTTAAAAAATTCCTTAAATAAAAAATAAGACTCGCTAGTCAGTAAGATATTGAAAT 1819
Qy 1805 atgatatgtccccctccggtgtctttgatcagagtgacatgtgcatcttttcaagggac 1864
Db 1820 ATGATATGCTGCCCTCCCGGTCTTTGATCAGGATGACATGTGCCATTTTTCAGAGGAC 1879
Qy 1865 gtgcagacaggtgctgcaattctagattacttctctgaaacatgctgtttggga 1924
Db 1880 GTGCAGACAGGCTGGCATTAGATTACTTTCTTACTCTGAAACATGGCTGTTTGGGA 1939
Qy 1925 gtgcgggattcaaaagtggtcccaacgctgcccactactgcaaatggcagtttttaacttta 1984
Db 1940 GTGCGGATTCAAAAGTGTGTCACCGCTGCCCCCTACTGCAATGGCAGTTTAAATCTTAA 1999
Qy 1985 tcttttgctctcagatggttgcaattgcatcttaacaaataatggtcagtcctc 2044
Db 2000 TCTTTTGGCTTCTCCAGATGGTTCGAATGATCTTAAACCAATATGTCAGTCTTCATC 2059
Qy 2045 tctgtctgtctcataggtgccaacctgtgtgtttaaagaaggaagctttgtacettta 2104
Db 2060 TCTGTCTGTCTCATAGTGTGCCACCTTGTGTGTGTTAAAGAAGGAAGCTTTGTACCTTAA 2119
Qy 2105 gagtgtaggtgaataatgaatggtgttgagtgacatgagaaacagcatgattcttg 2164
Db 2120 GAGTGTAGGTGAATGAATGCTTGGAGTGCATCTGAGAACGATATGATTTTGTG 2179
Qy 2165 ctttggggaataaataatgactgataatgtggtggtgtgttatttgagaagaataatca 2224
Db 2180 CTTTGGGGAATAAATAATGATGCTATCAAAATTTGGTGGGTGCTGATTTTGAAGAATAATCA 2239
Qy 2225 ttgcttatgcaaatggagctgaatttgataaaaaaaccaataacagctatgaagtgcg 2284
Db 2240 TTGCTTATGTCAATGGAGCTGAATTTGATAAAACCCCAATACAGCTATGAGTGTGCTG 2299

Qy 2285 ggcaagttacttt 2320
|||||
Db 2300 GGCAAGTTTACTTTTGTGATGTTTCTCCCACT 2335
RESULT 5
AF462268 2394 bp mRNA linear PRI 07-FEB-2002
LOCUS Homo sapiens UDP-glucuronosyltransferase UGT1A8*1 (UGT1A8) mRNA,
DEFINITION UGT1A8*1 allele, complete cds.
ACCESSION AF462268
VERSION AF462268.1 GI:18568218
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 2394)
AUTHORS Huang,Y.-H., Nguyen,N., Geske,D. and Tukey,R.
TITLE Identification and functional characterization of UGT1A8*1,
UGT1A8*2 and UGT1A8*3
JOURNAL Unpublished
AUTHORS 2 (bases 1 to 2394)
TITLE Direct Submission
JOURNAL Submitted (21-DEC-2001) Pharmacology, UCSD, 9500 Gilman Drive, La
Jolla, CA 92093, USA
FEATURES
Location/Qualifiers
1..2394
/organism="Homo sapiens"
/db_xref="taxon:9606"
/tissue_type="colon"
1..2394
/gene="UGT1A8"
/allele="UGT1A8*1"
64..1656
/gene="UGT1A8"
/codon_start=1
/product="UDP-glucuronosyltransferase UGT1A8*1"
/protein_id="AA175964.1"
/db_xref="GI:18568219"
/translation="MARTGWTSPICLVSLLLTCGPAEAGKLLVVPMDGSHWFTMQSV
VEKLLIRGHEVVVVMEVSNQISLNCVTYKTSYTLLEDLDFEDFADAQKQAV
RSLSFLSSNGFNLFTHQCRSLFNDRKLVEYLFKESSDAVFLDPFDACGLIVAKY
FSLPSVFEARGIACHLEGAOCAPLSPVPRILLGFSDAATFKERVRNHIHLEHL
FCQYSKNALEIAEILQTPVTAYDLYSHTSILLRDFVLDYPKPVPNNIMFTGGIN
CHQKPLPMEFEAVINAGSEHGVIVFSIGSMVSEIPEKKAMALADALGKIPQVLMRY
TGTRESLNANTILYKWLQNDLLGHPMTAFITHAGSHGVYESICNGYVPMVMPLEF
DOMDNAMKMETGAGVTNLVLEMTSDEENALKAVIDKXSKENIMRSLSHKDRPVE
PDLAVFWVEFWMRHKGAPHLPAADHLTWYQVHSLDVIQVLLAVLLVAFITFKCA
YGYRCLGKGRVKKHAKSKTH"
BASE COUNT 624 a 513 c 555 g 702 t
ORIGIN
Query Match 95.1%; Score 2207.4; DB 9; Length 2394;
Best Local Similarity 97.3%; Pred. No. 0;
Matches 2255; Conservative 0; Mismatches 61; Indels 1; Gaps 1;
Qy 5 gctgcagttctctgagtggttcacaggggtggaccagccccctctctctctgtgtgtgc 64
Db 50 GCTGCAGTTCTCTCATGGCTGCACAGGTTGGACAGCCCATTCCTATGTGTTTCTC 109
Qy 65 tgcctgtaacctgtgcttgcgagcagcaggaagctactgtagtgcctatgagtgagga 124
Db 110 TGCTCTGACCTGTGGCTTTGCTGAGCGAGGAGCTGCTGTGTAGTGCCTATGTGTTCTC 169
Qy 125 gccactggttcaacatgagtgctggtggagaaactattctcagggggcagtgaggtgg 184
Db 170 GTCACTGGTTCAACATGCAGTCGGTGGTGGAGAACTTATCCTCAGGGGCATGAGTGG 229
Qy 185 tigtgtatgcatgcagaggtgagttggaactggaagatcacctgaattgcacagtgaaga 244
|||||

Matches 2320; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy 1 ctcaagctgagttctctgagcttgacaggggtggaccagcccccctctctatgtgtg 60
Db 1 CTCAGCTGCAGTTCCTGATGGCTTGCACAGGGTGGACAGGCCCTTCTCTATGTGTG 60

Qy 61 tgcctctctgacacttgcttgcgagagcaggaagactactagtgagcccatgat 120
Db 61 TGTCCTCTGCTGACCTGTGCTGCTTGGCCAGCAGGAAGCTACTGTAGTCCCATGGAT 120

Qy 121 gggagccactgttccaccatgagtcggtggtggagaaactcatctcaggggcatgag 180
Db 121 GGGAGCCACTGTTCCACCATAGGTGCGTGGTGGAGAACTCATCTCAGGGGGCATGAG 180

Qy 181 gtggttgtagtcagcagagtgagttggaactgggaagatcaactgaattgcaagtg 240
Db 181 GTGTTGTGTAGTCATGCCAGAGGTGAGTTGGCAACTGGGAAGATCACTGAATTCACAGTG 240

Qy 241 aagacttattcaacttcataccctggagatctggaccgggagttcaaggcttttgc 300
Db 241 AAGACTTATTCAACTTCATATAGCCTGGAGGATCTGGACGGGAGTTCAGGCTTTTGGC 300

Qy 301 catgctcaatgaaagcacaaagtacgaagtatatattctctatttaagtggttcaaat 360
Db 301 CATGCTCAATGGAAGACRAAGTACGAAGTATATTTCTCTAATGGGTTCATAAAT 360

Qy 361 gacattttgacttatttttttcaaatgacaggagttgtttaagacaaaaaattagta 420
Db 361 GACATTTTGTGACTTATTTTCAAAATTCAGAGAGTTGTGTTAAAGACAAAAAATAGTA 420

Qy 421 gaatactaaaggagagttctttgatcaggtttctcgatcccttttgataactgtgc 480
Db 421 GAAATCTTAAGGAGAGTCTTTTGATGCAGTGTCTCGATCCCTTTTGATAACTGTGGC 480

Qy 481 ttaattgtgcaaatattctccctccctccctcggtcttcggcaggggaataactttgc 540
Db 481 TTAATTTGTGGCAAAATTTCTCCCTCCCTCCCTGCTGCTGCGCAGGGGAATACTTTGC 540

Qy 541 cactatctgaagaagtgacagtgccctctctctctctctctctctctctctctctctc 600
Db 541 CACTATCTTGAAGAGGTGCACAGTGCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTC 600

Qy 601 ttagggtctcagatgcacttcaagagagagatgcaggaacacatcatgcactgtg 660
Db 601 TTAGSGTCTCAGATGCCATGACTTTCAAGGAGAGATACGGAACACATCATCATGACTTG 660

Qy 661 gaggaacatttatatgcacogtttttcaaaaatgcctagaaatagcctctgaaatt 720
Db 661 GAGGAACATTTATATGCGACCGTTTTTCAAAAATGCCCTAGAAATAGCCTCTGAAAT 720

Qy 721 ctccaaacactgttacggagatgatctctacagccacacatcaattgtgtgcga 780
Db 721 CTCCAAACACCTGTACGGAGTATGATCTCTACAGCCACACATCAATTTGGTTGGCGA 780

Qy 781 acggactttgttttgactatccaaacccgtgatgcccaacatgatctctcattgtgtg 840
Db 781 ACGGACTTTGTTTGGACTATCCAAACCCGTGATGCCAACATGATCTTCAATTTGGTGT 840

Qy 841 atcaactgcactcggaaagccgttgctatggaaattggaagcctacattatgtctct 900
Db 841 ATCAACTGCGCATCAGGAAAGCCGTGTGCTATGGAATTTGAAGCCCTACATTAATTTCT 900

Qy 901 gsgaacaatggaattgtgtttctcttgggatcaatggtctcagaattccagagaag 960
Db 901 GGAGACATGGAATTTGTTTCTCTTTGGGATCAATGGTCTCAGAAATTTCCAGAGAG 960

Qy 961 aaagctatgcaattgtctgtatgctttgggcaaaatccctcagacagctcgtggcggtac 1020
Db 961 AAAGCTATGGCAATTTGCTGTGTTGGCGCAAAATCCCTCAGACAGTCCCTGTGGCGGTAC 1020

Qy 1021 actggaacccagcattcgaattctggaacacacagatattgttaagtgcctaccacaa 1080
Db 1021 ACTGGAAACCCGACCATGCAATCTTGGGAACACACAGTACTTGTAAAGTGGCTACCCCAA 1080
```

```
Qy 1081 aacgactgtctgtgcacccgatgaccogtgccttttatcaaccactgctgggtcccatgg 1140
Db 1081 AACGATCTGCTTGTGTACCCGATGACCCGTGCTTTTATCACCCATGCTGGTCCCATGGT 1140

Qy 1141 gtttatgaaacatacgcaatggcggttcccatgggtgatgatcccttgttgggtgatcag 1200
Db 1141 GTTTATGAAACATATGCAATGGCGTTCCCATGGTGTATGATGCCCTTTTGTGGTGTACAG 1200

Qy 1201 atggacaatgcaaaagcgcatggagactaaggagagctggagtgacccctgaattgtctggaa 1260
Db 1201 ATGGACAATGCAAAAGCGCATGAGAGACTAAGGGAGCTGGAGTGAACCTTGAATGTTCTGGA 1260

Qy 1261 atgaactctgaagatttagaaaaatgctctaaagaagcagtcataatgacaaagtacaag 1320
Db 1261 ATGACTCTGAAGATTTAGAAAAATGCTCTAAAGAGAGTATCAATGACAAAAGTTTCAAG 1320

Qy 1321 gagaacatcatgcccctctccagccttcaagaagcagcccgctggagccgctggagacctg 1380
Db 1321 GAGAACATCATGCGCCTCTCCAGCCTTCACAAAGGACCGCGCTGGAGCCGCTGGACCTG 1380

Qy 1381 gccgtgtctgggtgagtttgtgatagcacaagggcgccgaacctgcgcccgcga 1440
Db 1381 GCCGTGTTCTGGGTGAGTGTGTGATGAGGCACAAGGGCGCCACACCTGCGCCCGCGCA 1440

Qy 1441 gccacgacctcaactggtaccagttaccattccttggacgtgatgtgttctctctggcc 1500
Db 1441 GCCACGACCTCACCTGGTACCAGTACCATTCTTGGACGTGATTTGGTTCTCTCTGGCC 1500

Qy 1501 gtcgtctgacagtgaccttcacatttaaatgttgttcttgcctacacacgaagtc 1560
Db 1501 GTGCTGCTGACAGTGGCCTTCATCACCTTTAAATGTTGTCTTATGCTTATGGCTACCGGAATGC 1560

Qy 1561 ttggggaaaaagggcgagttaagaaagcccaaaatccaaagccattggagaagtgggt 1620
Db 1561 TTGGGAAAAAAGGGCGAGTTAAGAAAGCCACAAAATCCAAAGCCATGAGAAGTGGGT 1620

Qy 1621 gggaaataagtaaaattttgaaccattccctcagtcatttccaaacttgaaacagaatc 1680
Db 1621 GGGAAATAAGTAAATTTTGAACCATTTCCCTAGTCAATTTCCAAACTTTGAAACAGAAATC 1680

Qy 1681 agtgttaattcattttatttcttataaggaaatactttgtgcataaataatcagccccag 1740
Db 1681 AGTGTAAATTCATTTATTTCTTATTAAGGAAATFACTTTGTCATAAATTAATFACGCCCCAG 1740

Qy 1741 agtgccttataaaattctcttaataaaaaataatagactcgctagtcagtaaaagatatit 1800
Db 1741 AGTGCCTTTAAAAATTTCTCTTAATAAAAAATAATFAGACTCGCTAGTCAGTAAGATATTT 1800

Qy 1801 gaatatgtatcgccccctccggtgtctttgtatcaggatgacatgtgcatttttcaga 1860
Db 1801 GAATATGTATCGTGCCCCCTCCGGTGTCTTTGATFACAGGATGACATGTGCCATTTTTCAGA 1860

Qy 1861 gacgtgcagacagctggcattcttagattactttcttactctgaacatcgccgtgtt 1920
Db 1861 GGACGTGCAGACAGCTGGCATCTTAGATTACTTTTCTTACTCTGAAACATGGCCTGTTT 1920

Qy 1921 gggagtcgggattcgaaggtgtccacogctgcctcactcgtcaaatgycagttttaat 1980
Db 1921 GGGAGTCGGGATTCGAAAGGTGTGCCACCGCTGCCCTACTGCAATGGCAGTTTTTAAT 1980

Qy 1981 ctatcttttgctctctcagatgtgtcaattgatccttaaccaataatggtcagtcct 2040
Db 1981 CTATCTTTTGGCTCTCGCAGATGGTTCGAATGTATGCTTTAACCAATAATGGTCAGTCT 2040

Qy 2041 catctctcctgctctcaatagtgccacctgtgtgttttaaagaagggaagctttgtacc 2100
Db 2041 CATCTCTCTCTGCTTCATAGGTGCCACCTTGTGTGTGTTTAAAGAGGGAAGCTTTGTACC 2100

Qy 2101 tttaagtgtagtggaatgaatgaatgagcttgaggtgactgagaacacagcatatattt 2160
Db 2101 TTTAGAGTGTAGGTGAATGAATGAATGAGTGGCTTGGAGTGCATGACAGACGATATGATTT 2160
```

Qy	2161	cttgcttctgggaaaaagaatgatcgtcatgaatactggtggtggcgatttttgcagagata	2220
Db	2161	CTTGCCTTTGGGAAAAAGAAATCATCTATGAATTCGTGGTGCGTATTGTAAGAAGATA	2220
Qy	2221	atattgcc-tatgtcaaatggagctgaatttgtataaaaccccaaaaatacacagctatgaagt	2280
Db	2221	ATCAATTCCTTATGTCAAATGGAGCTGAATTTGTATATAAACCCCAAAATACAGCTATCAACT	2280
Qy	2281	gcaggcgaagttaacttttttctgatgttttcctacaact	2320
Db	2281	GCTGGCAGATTACTTTTTTCTGATGTTCTCTACAACCT	2320
RESULT	3		
LOCUS	AF056188		
DEFINITION	Homo sapiens UDP-glucuronosyltransferase 1A9 (UGT1A9) mRNA,	linear	PRI 13-AUG-1998
ACCESSION	AF056188	complete cds.	
VERSION	AF056188.1	GI:3414797	
KEYWORDS	.		
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Eutheraia; Primates; Catarrhini; Hominidae; Homo.		
TITLE	1 (bases 1 to 2320)		
JOURNAL	Ciotti,M., Potter,C. and Owens,I.S.		
AUTHORS	Human phenol metabolizing UDP-glucuronosyltransferase		
TITLE	Unpublished		
JOURNAL	2 (bases 1 to 2320)		
AUTHORS	Ciotti,M., Potter,C. and Owens,I.S.		
TITLE	Direct Submission		
JOURNAL	Submitted (30-MAR-1998) Heritable Disorders Branch, NICHHD, 9000		
REFERENCE	Rockville Pike, Bethesda, MD 20892, USA		
AUTHORS	3 (bases 1 to 2320)		
TITLE	Ciotti,M., Potter,C. and Owens,I.S.		
JOURNAL	Direct Submission		
REMARK	Submitted (12-AUG-1998) Heritable Disorders Branch, NICHHD, 9000		
COMMENT	Rockville Pike, Bethesda, MD 20892, USA		
FEATURES	Sequence update by submitter		
	On Aug 13, 1998 this sequence version replaced gi:3025895.		
	Location/Qualifiers		
source	1..2320		
	/organism="Homo sapiens"		
	/db_xref="taxon:9606"		
	/chromosome="2"		
	/map_2q37"		
gene	/tissue_type="liver"		
	1..2320		
CDS	/gene="UGT1A9"		
	19..1611		
	/gene="UGT1A9"		
	/codon_start=1		
	/product="UDP-glucuronosyltransferase 1A9"		
	/protein_id="AAC31425.1"		
	/db_xref="gi:3025896"		
	/translation="MACTGWTSPLCVCLLLTGFAEAGKLLVPMDGSHNTRSV		
	VEKLILRGHEVVVMWESLQGRSLNCTVKYTSYTLEDLDRFEKFAFAHQWAQV		
	RSIYSLLMGVSIVFDLPFTSNCRSLFKDKLVLYLKSSFOAVFLDPFNCGLIVAKY		
	FSLPSVPANGCILCHVLEGSCAGAPLSYVPRILLGFSDMTFKERVNRNHMLSEHL		
	LCHREFKNALITASEILQTPTEVDLSHSIWLLRTDFDVPKPVPNNPIFIGIN		
	CHQRPALPMEFEAYINASGEHGIVVFSLGMSSEIPEKKAMIAADALGKIPTQVLWRY		
	TGTRNSLNMMNTILVKMFPNDLLGHMPMTAFITHAGSHGYEVICNGVPMVMPFLVR		
	QDMDNKRMTKGATYLVKLEMTSEDENALKAVINDKSKENIMSLSSLHKDRPVE		
	PIDLAVWFVFWMRHKGAPHLRPAADHLTWYQVSHLDIGFLLAVALTVAFITFFCKA		
	YGRKCLDGKRVKKAHKSHTH"		
BASE COUNT	608 a 495 c 547 g 670 t		
ORIGIN			

|||||
Db 1861 GGAGGTGCAGACAGCTGGCATCTAGATTACTTTCTTACTCTGAACATGGCTGTTT 1920
Qy 1921 ggaggtgcagagattcaagggtggtccaccgctgcccctactgcaaaatggagtttaaat 1980
Db 1921 GGAGGTGCAGAGGATTCAAAAGTGGTCCACCGCTGCCCTACTGCAATGGGAGTTTAAT 1980
Qy 1981 ctatctcttggctctgagatggttgcattgacatctcttaaccataatggtcagtcct 2040
Db 1981 CTTATCTTTGGCTTCTGCAGATGGTTCGAATTCCTTAACCAATAATGTCAGTCT 2040
Qy 2041 catctctgctgcttcattaggtgcacactctctgttataaagaaggaagctttgac 2100
Db 2041 CATCTCTGCTCTTCATAGGTGCCACCTGTGTGTTTAAAGAGGAGCTTTGTACC 2100
Qy 2101 tttagagtgtagtgaaatgaatgaatggcttgagtgactgagacagcatatgatt 2160
Db 2101 TTTAGAGCTAGGTGAATGAATGAATGGCTTGGAGTGCACCTGAGAACAGCATATGATT 2160
Qy 2161 ctgcttgggaaagaagaatgatctatgaattggtggtgtatttgaagata 2220
Db 2161 CTTCCTTTGGGAAAAGAAATGATGCTATGAATTTGGTGGTGGTGTATTTGAGAAGATA 2220
Qy 2221 atcattgcttattgcaaatggagctgaatttgataaaaaaccccaatacagctatgaat 2280
Db 2221 ATCATGCTTATGCTCAATGAGCTGAATTTGATAAAAAACCCAAATACAGCTATGAAT 2280
Qy 2281 gctggagcaagtttactttttctgattgtttctcacaact 2320
Db 2281 CTGGGCAAGTTTACTTTTTTCTGATGTTTCTCTACAACT 2320
RESULT 2
AX286764
LOCUS AX286764 2320 bp DNA linear PAT 21-NOV-2001
DEFINITION Sequence 3 from Patent WO0180896.
ACCESSION AX286764
VERSION AX286764.1 GI:17048799
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (sites)
AUTHORS Ratalin,M.J., Innocenti,F. and Iyer,L.
TITLE Flavopiridol drug combinations and methods with reduced side effects
JOURNAL Patent: WO 0180896-A 3 01-NOV-2001;
ARCH DEVELOPMENT CORPORATION (US)
FEATURES
source 1..2320
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
19..1611
/note="unnamed protein product"
/codon_start=-1
/protein_id="CAD12275.1"
/db_xref="GI:17048800"
/translation="MACTGWTSPLCVLLTCGFABAGKLLVPMDSHWFMRVS
VEKILRGHEVVMPEVSWQLGRSLNCTVKVYSTYLEDLDRFKAFHAHQWKAQV
RSIYLLMGSYNDIFDLFFSNRSLFKDKLVEYLKESDFDAFLDPFNCGLIVAKY
FSLPSVVFARGILCHVLEEGACPAKLSVVPRTLLIGESDAMTFKERVNRHIMHEHL
LCHREFKNALEIASEILQTPVTEVDLYSHFSIWLRTDPLVDYKPKVPMNMFIGIN
CHQKPLPMFEAFYINASGEHGIUVFSLGSMSEIPEKKAMAIADALGKIPQVLWRY
TGTRPSLIANNITLVKWLFDNDLGHWPTRAFITIHAGSHGVIESICNGVPMWMLFPG
DOMNARKMETKAGVTLNVLEMTSEDLNALKAVINDKSYENIKMLSLHHRPVE
PLDLAVTFWFEVVRHKGAPHLRPAADHLTWQYVSHLDVIGFLAVLTVAFITFKCA
YGYRKLGLGKRGVRKHAHKSHTH"
BASE COUNT 608 a 495 c 547 g 670 t
ORIGIN

Best Local Similarity 100.0%; Pred. No. 0;
Matches 2320; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 ctaactgacattctctgactgctgacaggtgagagcagccctctctctatggtg 60
Db 1 CTGAGTGCAGGTTCTCTGATGGCTTGACAGGTTGGACAGGCTTCCCTCTATGTTG 60
Qy 61 tgtctgctgacacttggtcttgcagaggtgaggtgaggtgaggtgaggtgaggtgaggt 120
Db 61 TGTCTGCTGCTGACCTTGCTTGGCTTGGCAGGCGAGGGAAGCTACTGCTAGTGCCTAT 120
Qy 121 ggagagccactggttcacacagaggtcggtggtgagagaaactcatctcaggggcatgag 180
Db 121 GAGAGCCACTGGTTCACCATGAGGTGCGTGGTGGAGAACTCATCTCAGGGGCGATGAG 180
Qy 181 gttggttagtcaatgacagaggtgaggtgaggtgaggtgaggtgaggtgaggtgaggtgag 240
Db 181 GTGGTTGTAGTCACTCCAGAGAGGTGAGTTGGCAACTGGGAAGATCACTGAATGGCACAGTG 240
Qy 241 aagacttattcaactctatatacccttgaggtgaggtgaggtgaggtgaggtgaggtgaggtgag 300
Db 241 AAGACTTATTCAACTTCATATACCTGGAGGATCTGGACCGGGAGTTCAAGGCTTTTGGCC 300
Qy 301 catgctcaatggaaagcacaagtaoagagatataattctctatataatgggttatacaat 360
Db 301 CATGCTCAATGGAAAGCACAAGTACGAAGTATATATTTCTCTATTAAATGGGTTCATACAAT 360
Qy 361 gacattttgacttattttttcacaattgaggtggtttttaaagacacaaaaattagta 420
Db 361 GACATTTTGTGACTTATTTTTCAAATTCAGAGGATTTGTTTAAAGACAAAAATTAGTA 420
Qy 421 gaatactaaaggaggtcttcttgatgaggtggttctctgactctctctgactctctgactctctc 480
Db 421 GAATACTTAAAGAGAGAGTCTTTTGTATGACAGTGTTCATGATCCTTTTGTATGAATCTGGC 480
Qy 481 ttaattgttgcaaatattttctccctccctccctggtggtctctccagggaggaatacttgc 540
Db 481 TTAATTTTGGCAAAATATTTCTCCCTCCCTCCCTGCTGCTTCCGCCAGGGAATACTTTGC 540
Qy 541 cactatcttgaaagaggtgacaggtgacctgctctctctctctctctctctctctctctctc 600
Db 541 CACTATCTTGAAGAGAGGTGGACAGTGCCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTC 600
Qy 601 ttagggttctcagatgcatgacttcaaggagaggtgaggtgaggtgaggtgaggtgaggtgag 660
Db 601 TTAGGGTCTCAGATGCCATGACTTTTCAAGGAGAGAGTACGGAACCAACCATCATGCACTTG 660
Qy 661 gaggaacatttattatgccaccttttttcaaaaatgccctagaaatagcctctgaaatt 720
Db 661 GAGAACATTATTATGCCACCGTTTTCAAAAATGCCCTAGAAATAGCCTCTGAAATTT 720
Qy 721 ctcaaacacctgttcaggagatgatctctacagccacacacatcaattggttgggtgaga 780
Db 721 CTCCAAAACACCTGTTCAGAGATGATGATCTCTACAGCCACACATCAATTTGGTGTGGCA 780
Qy 781 acggaactgttttgactatcccaaacctgtatgcccaactgacatgactctctctctctctc 840
Db 781 ACGACTTTGTTTGGACTATCCCAACCCGTGATGCCCAACATGATCTCTCATTTGGTGGT 840
Qy 841 atcaactgcatcagggaagcoggttgcctatgggaatttgaagcctacattatgcttct 900
Db 841 ATCAACTGCCCATCAGGAAAGCCGCTTGCCTATGGAATTTGAAGCCTACATTAATGCTTCT 900
Qy 901 ggagaacatggaattggttttc 960
Db 901 GGAGACATGGAATTTGTTGGTTTCTCTTTGGGATCAATGGTCTCAGAAATTCAGAGAG 960
Qy 961 aaagctatggcaattgctgattggttgggcaaaatccctcagacagctctctctctctctct 1020
Db 961 AAAGCTATGGCAATTCCTGATGTTTGGGCAAAATCCCTCAGACAGTCTCTGTGGCGGTAC 1020
Qy 1021 actggaaccgacctgaatctctggaacacacagatacttctgtaagtggctaccccaa 1080
Db 1021 ACTGGAACCGACCTGAATCTCTGGAACACACAGATACTTCTGTAAGTGGCTACCCCAA 1080

Query Match 100.0%; Score 2320; DB 6; Length 2320;


```
/codon_start=1
/protein_id="CAD12274.1"
/db_xref="GI:17048798"
/translation="MACTGWTSPILCYCLLLTLCGFAEAGKLLVPMDSHWFTMRVY
VERLLLRHEVVVMPVSVNQLGSLNCTVKTYSYTLLEDDLRFEKFAHQAQKAY
RSIYLLKMSYNDIFDLFFNCRSLFKKLLVETLKESSDAVFLDPDNCGLLVAKY
FSFVSYPFARGLILYLEBQAGPAPLSYVPRILLGFSDAMTFKRVNRHHIMHLEHL
LCIRFPKNALEIASIILQTVTEYDILYSHTSWLLRTDFVLDYPKRVPNNMIFIGGN
CHQKPLPMEFFAYINASGEHGVFVSLGSMVSEIPEKKAMAIADALGIKIQTVLWRY
TGRPSLANNTILVKKLPQNDLLGHPTRAFITHAGSHGVYESICNGVPMVMPPLFG
DOMDNAKMETKAGVTLNLEMTSEDLENALKAVINDKSYKENIMRSLSLHKKRPVE
PLDLAVFWFEFVMRHKGAPHLRPAADHLTWQYHSLDIVIGLLAVLVLTVAITEKCA
VGYRKCLGKKGKGVKKAHKSHTH"
```

BASE COUNT 608 a 495 c 547 g 570 t

ORIGIN

```
Query Match 100.0%; Score 2320; DB 6; Length 2320;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2320; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ctgagtcgagtcctctgagtgctgcacagggtggaacagcccccctctctctatgtg 60
DB 1 CTCAGTCGAGTCTCTCATGCGTTGACAGGGTGGACAGCCCTCTCTCATGTGTG 60

QY 61 tgcctgctgacatgctggcttgcagagggaggaagcactgctgtagtgcacatggat 120
DB 61 TGCTGCTGCTGACCTGTGGCTTTCAGAGGCGAGGGAAGCTACTGTGTAGTGCCTATGGAT 120

QY 121 gggagccactggttcacatgaggtcggtggtggaagaaactcatctcaggggagcatgag 180
DB 121 GGGAGCCACTGGTTCACCATGAGTGGTGTGGGAGAACTCATCTCTCAGGGGCGATGAG 180

QY 181 gtaggttagtcagtcagaggtgaggtggcaactgggaagatcactgaaatggacagtg 240
DB 181 GTGGTGTAGTCATGCGCAGAGGTGAGTTGGCACTGGGAAGATCATCTGAATGGCACAGTG 240

QY 241 aagaacttatcaacttcataccctggagagatctggacgggaggttcaaggcttttgc 300
DB 241 AAGACTTATCAACTTCATATACCTGGAGGATCTGGACGGGAGTTCAGAGCTTTTGGC 300

QY 301 catgctcaatggaaagcaagtaagaagtatatattctctctatttaattgggttcatacaat 360
DB 301 CATGCTCAATGGAAAGCACAAAGTACGAAGTATATATCTCTATTAATGGGTTCATACAAT 360

QY 361 gacattttgacttattttttcaaatgcagaggtgtgttttaaaagacaaaataatagta 420
DB 361 GACATTTTGTGACTTATTTTTTCAAAATTCAGAGGAGTTTGTTTAAAGACAAAAATTAGTA 420

QY 421 gaatacttaagagagagattctttgatcagtggttctcgatcctcttttgataactgtggc 480
DB 421 GAATACTTAAAGGAGAGTCTTTTGTATGACAGTGTCTCGATGCTTTTGTATTAACGTGGGC 480

QY 481 ttaattgttgccaataattctccctccctccctgggtgtcttcggcagggggaataactttgc 540
DB 481 TTAATGTGTGCAAAATATTTCTCCCTCCCTCCCTGGTGTCTTCGCCAGGGGAATACTTTGC 540

QY 541 cactacttgagagagtgacagtgccctgctctctctctctctctctctctctctctctc 600
DB 541 CACTATCTTGAAGAGGTGCACAGTGCCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTC 600

QY 601 tttaggttctcagatgcctatgactttcaagagagagtagtcggaacacatcagcacttg 660
DB 601 TTAGGTTCTCAGATGCATGACTTTTCAAGGAGAGAGTAGCGGAACACCATCATGCACTTG 660

QY 661 gaggaacatttattatgccacggttttttcaaaatgccctagaaatagcctctgaatt 720
DB 661 GAGGAACATTTATTTATGCCACCGTTTTTCAAAAATGCCCCCTAGAAATAGCCCTCTGAATTT 720

QY 721 ctccaaacacctgttaccgagtagtatctctacagccacacatcatcaatttgggttgcga 780
DB 721 CTCCAAACACCTGTTACGGAGTAGTATCTCTACAGCCACACATCAATTTGGTTGTTGCCA 780
```

```
QY 781 acggacttgttttgagactatcccaaacccgctgatgcccaacatgatctctcattgtggt 840
DB 781 ACGGACTTGTTTTGGACTATCCCAAAACCCGCTGATGCCCAACATGATCTTCATTTGGTGGT 840

QY 841 atcaactgcacatcaggaagacgctgtgctatggaatttgaagcctacattaatgtctct 900
DB 841 ATCAACTGCCATCAGGAAGACCGTGTGCTATGGAATTTTGAAGCCTACATTAATGTGCTCT 900

QY 901 ggagaaactggaaattgtggttttcttcttggatgtaaatgctcagaaattccagagaag 960
DB 901 GGAGAACA"CGAATTGTGGTTCCTCTTTGGATCAATGCTCTCAGAAATTCAGAGAGAG 960

QY 961 aaagctatggcaattgctgactgttgggaaaaaacctcccaagcagtcctgtggcgggtac 1020
DB 961 AAAGCTATGGCAATTGCTGATGCTTTGGGCAAAATCCCTCAGACAGTCTCTGTGGCGGTAC 1020

QY 1021 actggaacccgaccatcgaaatcttgcgaacacacgatactgttaagtggctaccccaa 1080
DB 1021 ACTGGAACCCGACCATCGAATCTTGCGAACAACACGATACTTGTTAAGTGGCTACCCCAA 1080

QY 1081 aacgactgctgtgtcaacggatgacccgtgcctttatcaccatgctggttcccatggt 1140
DB 1081 AACGATGCTGTGTGTACCCGATGACCCGTGCTTTATCACCCTATGCTGGTTCCTCATGGT 1140

QY 1141 gttatgaagcatatgcaatggcgttcccatgggtgatgatgcttctgttgggtgatcag 1200
DB 1141 GTTATGAAGCATATGCAATGGCTTCCCATGCTGATGATGATGCCCTTGTTTGGTGTATCAG 1200

QY 1201 atggacaatgcaagcgcatgagactaaaggagactggagactgagactgagactgagactgagaa 1260
DB 1201 ATGGACAA"GGCAAGCGCATGAGACTAAGGGAGCTGGAGTGAACCTCGAATGTTCTTGGAA 1260

QY 1261 atgacttcgaagatttagaaaaatgctctaaaagcagtcacatgaacaaagtacaaag 1320
DB 1261 ATGACTTC"GAAGATTTAGAAAATGCTCTAAAAGCAGTCAATCAATGACAAAAGTTACAAG 1320

QY 1321 gagaacatcatgcccctctccagccttcacaagagacccgctggagccgctggagccg 1380
DB 1321 GAGAACATCATGCGCTCTCCAGCCTTCCAAAGACCGCCCGGTGGAGCCGCTGGACCTG 1380

QY 1381 gcccgttctgggtgaggttctgtagggcacaagggcgcccaacacctgcgcccgcga 1440
DB 1381 GCGGTGTTCTGGGTGAGGTTTGTGATGAGGCACAAGGGCGCGCACACCTGCGCCCGCA 1440

QY 1441 gcccagcctccacctggtaccagtaaccattcctctggagctgattggtttcctctctggcc 1500
DB 1441 GCCCAGCACTCACCTGGTACCAGTACCAATTCCTTGGACGTGATTTGGTTCCTCTTGGCC 1500

QY 1501 gctgtgctgaacagtgcccttcacaccttttaaatgttgttcttatggctaccggaaatgc 1560
DB 1501 GTCGTGCTGACAGTGGCCTTTCATCACCTTTAAATGTTGTGCTTATGGCTACCGGAATGC 1560

QY 1561 ttggggaaaaagggcgagttgaagaaagcccaaaaatccaaagaccctattgagaagtgggt 1620
DB 1561 TTGGGAAAAAAGGGCGAGTTAAGAAAGCCCAAAATCCAAAGACCCATTTGAGAAGTGGGT 1620

QY 1621 gggaataagtaaaattttgaacattccctgaacttcccaacttgaaacagaatc 1680
DB 1621 GGAATAAGTAAATTTTGAACCATTCCTTAGCTATTTCCAAACTTTGAAACAGCAATC 1680

QY 1681 agtgttaattcattttatttcttatttaagggaataactttgataaatttaacagccccag 1740
DB 1681 AGTGTAAATTCATTTTATTTCTTATTAAAGGAAATACCTTTTGCATAAATTAATACGCCCCAG 1740

QY 1741 agtgccttaaaaaatctcttaataaaaaataatagactcgctagctagtaagaatatit 1800
DB 1741 AGTGCCTTAAAAAATCTCTTAATAAAAAATAATAGACTCGCTAGCTAGTAAGATATTT 1800

QY 1801 gaatgtatgctgccccctccgggtgcttttggatcaggatgacatgtgccaatttttcaga 1860
DB 1801 GAATATGATCGTCCCCCTCCGGGTCTTTTGTGATCAGGATGACATGTGCCATTTTTCAGA 1860

QY 1861 ggaactgcagacaggtggcattcttagattacttttcttactctgaaacatggcctgttt 1920
```


RESULT	14	
BG469506		
LOCUS		
DEFINITION	829 bp mRNA linear EST 21-MAR-2001	
	602532818f1 NIH_MGC_15 Homo sapiens cDNA clone IMAGE:4660540 5',	
	mRNA sequence.	
ACCESSION	BG469506	
VERSION	BG469506.1	GI:13401781
KEYWORDS	EST.	
SOURCE	human.	
ORGANISM	Homo sapiens	
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.	
REFERENCE	1 (bases 1 to 829)	

[illegible]

```

ACCESSION AI831562
VERSION AI831562.1 GI:5452233
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 764)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncilogap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
JOURNAL Tumor Gene Index
COMMENT Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-re@mail.nih.gov
Tissue Procurement: Christopher Moskalko, M.D., Ph.D., Michael R.
Emmert-Buck, M.D., Ph.D.
cDNA Library Preparation: M. Bento Soares, Ph.D.
cDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Cloned through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html
Insert Length: 1491 Std Error: 0.00
Seq primer: -40UP from Gibco
High quality sequence stop: 454.
FEATURES
source
1..764
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="NCI-CGAP_Lu19"
/tissue_type="squamous cell carcinoma, poorly
differentiated (4 pooled tumors, including primary and
metastatic)"
/dev_stage="adult"
/lab_host="DH10B (phage-resistant)"
/notes="Organ: lung; Vector: p773D-Pac (Pharmacia) with a
modified polylinker; 1st strand cDNA was prepared from
pooled lung tumor tissue, and was then primed with a Not I
- oligo(dT) primer. Double-stranded cDNA was ligated to
Eco RI adaptors (Pharmacia), digested with Not I and
cloned into the Not I and Eco RI sites of the modified
p773 vector. Library went through one round of
normalization. Library constructed by Bento Soares and M.
Fatima Bonaldo."
BASE COUNT 242 a 156 c 133 g 232 t 1 others
ORIGIN
Query Match 28.9%; Score 669.4; DB 9; Length 764;
Best Local Similarity 96.8%; Pred. No. 2.5e-164;
Matches 693; Conservative 0; Mismatches 22; Indels 1; Gaps 1;
QY 1605 ccattgagaagtgggtgggaataagtgataattttgaaccattccctagtcatttccaa 1664
Db 755 CCCATGAAGAGTGGGTGGGAATAGGCTAAATTTTGAACCATCCCTACGTCATCTTCCAA 696
QY 1665 acttgaacaagaatcagtgtaattcattttatttatttatttatttatttatttatttgcata 1724
Db 695 CTTTGAACACAGCATCAAGTGGTAAATTCATCTTATTAAGAAAATATCTTTGCATA 636
QY 1725 aattaatcagcccgagtgcttttaaaaaattctcttaaaaaataaataaataaataaataaataa 1784
Db 635 AATTAATCAGCCCGAGAGTGC-TTAAANAATCTCTTAAANAATCTCTTAAANAATCTCTTAA 577
QY 1785 gtctgtaaaagatttgaattatgtatgtatgtatgtatgtatgtatgtatgtatgtatgtatgtatgaca 1844
Db 576 GTCACTACAGATATTTGAATATGATCGTCCGCCCTCCCGGTGCTCTTTGATCAGGATGACA 517
QY 1845 tttgcatttttccagagacgtgcagacagcgtgcattctcagatttcttcttcttcttcttcttctt 1904
Db 516 TGTGCCATTTTTCAGAGGACGTGCAGACAGCGCTGGCATCTCTAGATTACTTCTTCTACTCT 457
QY 1905 gaaacatggcctgttttggagtcgaggattcaaaagtggtccacacgctgcacctactgc 1964
Db 456 GAAACATGGCCTGTGGAGTGGCGGATTCAAAGGTGGTCCACCGCTGCCCTACTGC 397
QY 1965 aaatggcagttttaaattcttcttcttcttcttcttcttcttcttcttcttcttcttcttcttctt 2024
Db 396 AAATGGCAGTTTAAATCTATCTTTTGGCTTCGAGATGGTTGCAATGATGATCCTTAACC 337
QY 2025 aataatggctagtcctctctcttcttcttcttcttcttcttcttcttcttcttcttcttcttctt 2084
Db 336 AATAATGGT-AGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 277
QY 2085 aggaagacttcttcttcttcttcttcttcttcttcttcttcttcttcttcttcttcttcttcttctt 2144
Db 276 AGGAAGACTTGTACTCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 217
QY 2145 gaacagcatatgattcttcttcttcttcttcttcttcttcttcttcttcttcttcttcttcttctt 2204
Db 216 GAACAGCATATGATTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 157
QY 2205 tctattgagaagataaatacttcttcttcttcttcttcttcttcttcttcttcttcttcttcttcttctt 2264
Db 156 TCTATTTGAGAGATTAATCATCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 97
QY 2265 aatacagctatgaatgctgggcaagtcttcttcttcttcttcttcttcttcttcttcttcttcttctt 2320
Db 96 AATACAGCTATGAAGTGGTGGCAAGTTTACTTTTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 41
RESULT 13
BG424452
LOCUS 723 bp mRNA linear EST 14-MAR-2001
DEFINITION BG424452 Homo sapiens cDNA clone IMAGE:4586368 5',
mRNA sequence.
ACCESSION BG424452
VERSION BG424452.1 GI:13330958
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 723)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-re@mail.nih.gov
Tissue Procurement: DCTD/Drp
cDNA Library Preparation: Ling Hong/Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Cloned through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LCM1314 row: p column: 17
High quality sequence stop: 711.
FEATURES
source
1..723
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="NIH_MGC_14"
/tissue_type="renal cell adenocarcinoma"
/lab_host="DH10B (phage-resistant)"
/notes="Organ: kidney; Vector: pOTB7; Site_1: XhoI; Site_2:
EcoRI; cDNA made by oligo-dT priming. Directionally
cloned into EcoRI/XhoI sites using the following 5'
adaptor: GGCACGAG(G). Size-selected >500bp for average
insert size 1.8kb. Library constructed by Ling Hong in
the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies)."
BASE COUNT 202 a 160 c 170 g 191 t

```

```

Db 181 GGAGACTAAGGAGCTGAGTACCCTGAATGTTCTGGAAATGACTTCTGAAGATTAGA 240
Qy 1281 aaatgcttaaaagcagtcatacaaaagttacaagaggaacacatcagcgcctc 1340
Db 241 AAATGCTCTAAAGCAGTCAATCAATGACAAAGTTCAAGAGGAACATCATGCGCTCTC 300
Qy 1341 caagctcaagagacagcccggtgagacgcctgagacccgtgcgtctctggtgagtt 1400
Db 301 CAGCCTTCAAGAGACCCGCGGTGGAGCCGCTGGACCTGGCCGTGTTCTGGGTGGAGTT 360
Qy 1401 tgtgatgaggcacaagggcgccacacactgcccccgagcccgacccctcaactggta 1460
Db 361 TGTGATGAGGACAAAGGCGCCACACCTGCGCCCGCAGCCACGACCTCACTGTTA 420
Qy 1461 ccagtaaccattcctgagagtgattggttctctctgctgcccgtcgctgacagtgccct 1520
Db 421 CCAGTACCATTCCTTGGAGCTGATTTGGTTTCTCTTGGCCGTGCTGACAGTGCCCTT 480
Qy 1521 catcacctttaaattgtgtgcttatggctaccggaatgcttggggaaaaaaggcgagt 1580
Db 481 CATCACCTTAAATGTTGTGCTTATGGCTACCGGAAATGCTTGGGAAAAAGGCGAGT 540
Qy 1581 taagaagaccacaaatccaagacccattgagaagtggtggtggaataaagtaaaatttt 1640
Db 541 TAAGAAGCCCAAAATCAAGACCCATTGAGAAGTGGGTGGGAAATAAGGTAAATTTG 600
Qy 1641 gaaccattccct-agtcatttccaaacttgaaacagaaatcagtggttaattcatittat 1699
Db 601 GAACCATTCCTTAAGTCAATTCCTCAAACTTGAACAGCAATCATGTTCAATTCATCT--A 658
Qy 1700 tcttattaagaataacttgcataataataataataataataataataataataataatctc 1759
Db 659 TCTATTACGAAATCTTCGCTCATATT-ATCAGCCGAGAGTCTCTT-ACAAATCCTC 716
Qy 1760 ttaataaaaaataatagactcgcgtagtcagtaaatattt 1800
Db 717 TAAATAACAAATAATAGACTCGCTAGTCAGAACAGATCTCT 757

```

RESULT 11

```

LOCUS BG747143 697 bp mRNA linear EST 15-MAY-2001
DEFINITION 602704364F1 NIH_MGC_15 Homo sapiens cDNA clone IMAGE:4857831 5',
mRNA sequence.
ACCESSION BG747143
VERSION BG747143.1 GI:14057796
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 697)
NIH-MGC http://imgc.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Ling Hong/Rubin Laboratory
DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: NIH Intramural Sequencing Center
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LICM1710 row: o column: 16
High quality sequence stop: 684.
Location/Qualifiers
i. . 697
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4857831"
/clone_lib="NIH_MGC_15"

```

FEATURES

source

```

RESULT 12
LOCUS AI831562/c 764 bp mRNA linear EST 21-DEC-1999
DEFINITION wj39b01.x1 NCI_CGAP_Lu19 Homo sapiens cDNA clone IMAGE:2405161 3',
similar to gb-M57899 UDP-GLUCURONOSYLTRANSFERASE 1A PRECURSOR,
MICROSOMAL (HUMAN);, mRNA sequence.

```

```

/tissue_type="adenocarcinoma cell line"
/lab_host="DH10B (phage-resistant)"

```

```

/Note="Organ: colon; Vector: pOTB7; Site_1: XhoI; Site_2:
EcoRI; cDNA made by oligo-dT priming. Directionally
cloned into EcoRI/XhoI sites using the following 5'
adaptor: GGCAGAG(G). Size-selected >500bp for average
insert size 1.8kb. Library constructed by Ling Hong in
the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies)"

```

```

BASE COUNT 177 a 168 c 184 g 168 t
ORIGIN

```

```

Query Match 29.3%; Score 679.2; DB 10; Length 697;
Best Local Similarity 99.4%; Pred. No. 6.5e-167;
Matches 692; Conservative 0; Mismatches 3; Indels 1; Gaps 1;

```

```

Qy 910 ggaattgtggtttctcttctgtggatcaatggtctcagaattccagagagaagaagctatg 969
Db 2 GGAATTGTGTTTCCTTTGGGATCAATGGTCTCAGAAATTCAGAGAAAGAGCTATG 61
Qy 970 gcaattgctgattgttgggcaaatccctcagacagtcctgtggtggtacactggaacc 1029
Db 62 GCAATTGTGATGCTTGTGGCAAAATCCCTCAGACAGTCTCTGTGGCGGTACTACTGGAAC 121
Qy 1030 cgaccattcgaattctgcgaacacacgatacttcttaagtgcctaccaccaaacgactg 1089
Db 122 CGACCATCGAATCTTGGAAACAACAGCATACTTGTAAAGTGCTACCCCAAAACGATCTG 181
Qy 1090 ctgtgaccccgatgacccgtgctcttataccacatgctggttcccatggtgtttatgaa 1149
Db 182 CTGTGTCACCCGATGACCCGCTTATCACCATGCTGTGTCCTCATGTTGTTATGAA 241
Qy 1150 agcatatgcaatgcgttcccatggtgatgacctgttctgttgatcagatggacaat 1209
Db 242 AGCATATGCAATGGCTTCCCATGTTGATGATGCCCTTGTGTTGGTGTATCAGATGGACAAT 301
Qy 1210 gcaaaagcagatgagactaaaggagctggtgagcctgaactgttctggaatgaactct 1269
Db 302 GCAAAAGCCATGGAGACTAAGGGAGCTGGAGTACCCTGAATGTTCTGGAATGACTTCT 361
Qy 1270 gaagatttagaaaatgctctaaagcagtcacatgacaaagtacaaaggagagaacatc 1329
Db 362 GAAGATTTAGAAAATGCTCTAAAGACAGTCAATCAATGACAAAAGTTACAAAGAGAACATC 421
Qy 1330 atgagcctctccagccttcaagaagaccggtgagcagcctgagacccctggccgtgttc 1389
Db 422 ATGCCCTCTCCAGCCTTCAAGAGACCCGCGGTGGAGCCGCTGGACCTGGCCGCTGTTCT 481
Qy 1390 tgggtgagtttgtgatgaggcacaaggcgccgcccacactgcccccgagccccagac 1449
Db 482 TGGGTGGAGTTTGTGATGAGGCACAAGGGCGCCACACCTGCGCCCGCAGCCACGAG 541
Qy 1450 ctcaactgttaccactaccattccttggagtgattggttctctcttggcgcgtcgctg 1509
Db 542 CTCACCTGGTACCAGTACCATTCCTTGGACGTGATTTGTTTCTTGGCCGCTGTCGTCG 601
Qy 1510 acagtggccttcacactttaaattgttggcttatggctaccggaatgcttggggaaa 1569
Db 602 ACAGTGGGCTTCATCACCCTTTAAATGTTGTGCTTATGCTACCGGAAATGCTTGGGAAA 661
Qy 1570 aaaggcgagttgaagaagcccaaatccaagacc 1605
Db 662 AAAGGCGAGTTAAGAAG-CCACAAATCCAAAGACC 696

```

CDNA Library Preparation: Life Technologies, Inc.
 DNA Library Arrayed by: The I.M.A.G.E. Consortium
 DNA sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
 Plate: LLAM10401 row: n column: 18
 High quality sequence set: 763.

5
5
4
5
5
2

3
6
4
8
0
2

```

1: /root
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4514129"
/clone.lib="NIH_MGC_93"
/tissue_type="transitional
/lab_host="DH10B (phage-res
/notes="Organ: bladder; Vect
Site.2: SalI; Cloned undir
Average insert size 1.7 kb.
Length clones and cons
full-length clones and cons
Note: this is a NIH_MGC lib
263 a 154 c 198 q 266 t
BASE COUNT

```

```

/clone_lib="NIH_MGC_93"
/tissue_type="transitional cell papilloma, cell line"
/lab_host="DH10B (phage-resistant)"
/notes="Organ: bladder; Vector: pCMV-SPORT6; Site_1: NOT1;
Site_2: SalI; Cloned unidirectionally; oligo-dt primed.
Average insert size 1.7 kb. Library enriched for
full-length clones and constructed by Life Technologies.
Note: this is a NIH_MGC Library."
263 a 154 c 198 g 266 t

BASE COUNT
ORIGIN

Query Match 30.0%; Score 695.4; DB 10; Length 881;
Best Local Similarity 97.7%; Pred. No. 4e-171;
Matches 769; Conservative 0; Mismatches 11; Indels 7; Gaps 6;

```

	BASE COUNT	ORIGIN
1	0.000	0.000
2	0.000	0.000
3	0.000	0.000
4	0.000	0.000
5	0.000	0.000
6	0.000	0.000
7	0.000	0.000
8	0.000	0.000
9	0.000	0.000
10	0.000	0.000
11	0.000	0.000
12	0.000	0.000
13	0.000	0.000
14	0.000	0.000
15	0.000	0.000
16	0.000	0.000
17	0.000	0.000
18	0.000	0.000
19	0.000	0.000
20	0.000	0.000
21	0.000	0.000
22	0.000	0.000
23	0.000	0.000
24	0.000	0.000
25	0.000	0.000
26	0.000	0.000
27	0.000	0.000
28	0.000	0.000
29	0.000	0.000
30	0.000	0.000
31	0.000	0.000
32	0.000	0.000
33	0.000	0.000
34	0.000	0.000
35	0.000	0.000
36	0.000	0.000
37	0.000	0.000
38	0.000	0.000
39	0.000	0.000
40	0.000	0.000
41	0.000	0.000
42	0.000	0.000
43	0.000	0.000
44	0.000	0.000
45	0.000	0.000
46	0.000	0.000
47	0.000	0.000
48	0.000	0.000
49	0.000	0.000
50	0.000	0.000
51	0.000	0.000
52	0.000	0.000
53	0.000	0.000
54	0.000	0.000
55	0.000	0.000
56	0.000	0.000
57	0.000	0.000
58	0.000	0.000
59	0.000	0.000
60	0.000	0.000
61	0.000	0.000
62	0.000	0.000
63	0.000	0.000
64	0.000	0.000
65	0.000	0.000
66	0.000	0.000
67	0.000	0.000
68	0.000	0.000
69	0.000	0.000
70	0.000	0.000
71	0.000	0.000
72	0.000	0.000
73	0.000	0.000
74	0.000	0.000
75	0.000	0.000
76	0.000	0.000
77	0.000	0.000
78	0.000	0.000
79	0.000	0.000
80	0.000	0.000
81	0.000	0.000
82	0.000	0.000
83	0.000	0.000
84	0.000	0.000
85	0.000	0.000
86	0.000	0.000
87	0.000	0.000
88	0.000	0.000
89	0.000	0.000
90	0.000	0.000
91	0.000	0.000
92	0.000	0.000
93	0.000	0.000
94	0.000	0.000
95	0.000	0.000
96	0.000	0.000
97	0.000	0.000
98	0.000	0.000
99	0.000	0.000
100	0.000	0.000

BASE COUNT	282 a	268 c	246 g	209 t
------------	-------	-------	-------	-------

Query Mat
Best Loca
Matches

1041	QY	tcttgcgaacacacgatactgttaagtggctacccccaaaacgatactgtgtgtcacc	1100
1	Db	TCTTGGCAACAACACGATACCTGTTAAGTGGCTACCCAAAACGATCTGTTGTCACCC	60
1101	QY	gatgaaccgctgcctttatcacccatgctggtttcccatggtgtttatgaaagcatatgaa	1160
61	Db	GATGACCCGTGCCCTTATCACCCATGCTGGTGTCCCATGTGTTTATGAAAGCATATGCAA	120
1161	QY	tggcgttcccatggigtatgacccctgttggatgacagatggacaatgcaaaagcgcat	1220
121	Db	TGGCGTCCCATGGTGATGATGCCCTGTTGGTGATCATAGATGGACAAATCAAAAGCGAT	180
1221	QY	ggagactaaggagactggagtgaacctgaattgttcggaaatgactctgaagattaga	1280

QY 683 gtttttcaaaatgcctag--aaatgctctgaaatctcc-aaacacctgttaagg 739
 Db 660 GTTTTTCAAAATGCCCTAGGAATAGCCTCTGAAATNCTCCAAAACACCTGTACGG 719
 QY 740 agtatgatctacagcacacacataaatttggttggtgcaacgacacitgttttgact 799
 Db 720 AGTATGATCTCTACAGCACACATCAATCTGGCTCGTGGCAACGACGACTGTCTTGACT 779
 QY 800 atccaaaacccgtagcccaacatgatcttcattggtggtatcaac--tgccatcaggg 857
 Db 780 AT-CCAAACCCGATGATGCCACAAATGATCCTCTTGGGGGTATCCACCTGCCGCTACGG 838
 QY 858 aaagcgttgctatgaattga 881
 Db 839 AAAGCGTTGCCCTATGGATTGGA 862

RESULT 8
 AI307269/c 818 bp mRNA linear EST 08-APR-1999
 LOCUS tb24d12.x1 NCI-CGAP_Kid12 Homo sapiens cDNA clone IMAGE:2055287 3'
 DEFINITION similar to gb:457899 UDP-GLUCURONOSYLTRANSFERASE 1A PRECURSOR,
 MICROSMAL (HUMAN);, mRNA sequence.

ACCESSION AI307269
 VERSION AI307269.1 GI:4002458
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 818)
 NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
 National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
 Tumor Gene Index
 Unpublished (1997)
 CONTACT: Robert Strausberg, Ph.D.
 Email: cagaps-remail.nih.gov
 Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
 Emmert-Buck, M.D., Ph.D.
 CDNA Library Preparation: M. Bento Soares, Ph.D.
 CDNA Library Arrayed by: Greg Lennon, Ph.D.
 DNA Sequencing by: Washington University Genome Sequencing Center
 Clone Distribution: NCI-CGAP clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
www.bio.lnl.gov/bbrp/image/image.html
 Seq primer: -40UP from Gibco
 High quality sequence stop: 453.
 Location/Qualifiers
 1..818
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:2055287"
 /clone.lib="NCI-CGAP_Kid12"
 /tissue_type="2 pooled tumors (clear cell type)"
 /lab_host="DH10B"
 /note="Organ: kidney; Vector: p7f3D-Pac (Pharmacia) with
 a modified polylinker; Site_1: Not 1; Site_2: Eco RI;
 Plasmid DNA from the normalized library NCI-CGAP_Kid5 was
 prepared, and ss circles were made in vitro. Following HAP
 purification, this DNA was used as tracer in a subtractive
 hybridization reaction. The driver was PCR-amplified cDNAs
 from a pool of 5,000 clones made from the same library
 (cloneIDs 1323912-1325831, 1471368-1472903 and
 1492104-1493255). Subtraction by Bento Soares and M.
 Fatima Bonaldo."

BASE COUNT 247 a 177 c 134 g 258 t 2 others
 ORIGIN

Query Match 30.0%; Score 695.8; DB 9; Length 818;
 Best Local Similarity 95.9%; Pred No. 3e-171;
 Matches 745; Conservative 0; Mismatches 29; Indels 3; Gaps 3;

QY 1547 gctaccggaaatgcttggggaaaaagggcg-agttaagaagccccacaataccaagacc 1605
 Db 818 GCTACGAAAATGCTGGGGAAAAAAGGGCGAATTTAAGNAAGCCCNCAATCCCAAGAAC 759
 QY 1606 catgaagaagtggtgggaaataagtgtaaaatttgaacattccctagctcattccaa 1665
 Db 758 ATTTGGAAGTGGGTGGGAATAAGGTAAATTTGAACCAATTCGCCCAAGTCATTTCCCAA 699
 QY 1666 ctgaaaaacagaatca-gtgttaattcatttatttatttatttaagaagaataactttgca 1724
 Db 698 CTGAAAACAGAAATCAGGGGTAAAAATTCATTTTATTATTAAAGAAATACTTTGGCATA 639
 QY 1725 aattaatcagcccc-agagtgtctttaaataattctcttaataaaaaataatagactcgct 1783
 Db 638 AATTAAATCAGCCCCAAGAGTGCTTTTAAATAATCTCTTAAATAAAATAATAGACTCGCT 579
 QY 1784 agtcagtaagaatatttgaatatgtatgtgccccctccgggtgtcttttgatcagatgac 1843
 Db 578 AGTCAGTAAAGATATTTGAATATGTATGTCGCCCTCTGGTCTTTGATCAGATGAC 519
 QY 1844 atgtaccatttttcagagacgtgcagacagctggcattcctagattacttttcttactc 1903
 Db 518 ATGTGCATTTTTCAGAGGACGTGCAGACAGCGTGGCAATTTAGATTACTTTTCTTACTC 459
 QY 1904 tgaacatggcctgtttggggagtgcggtgattcaaaagtggtcccccacgcgtccctactg 1963
 Db 458 TGAACATATGCTGTTGGGAGTGGCGGATTCAAAAGGTGGTCCACGCTGCCCTACTG 399
 QY 1964 caaatggcagtttaacttcttcttcttcttcttcttcttcttcttcttcttcttcttctt 2023
 Db 398 CAATGGCAGTTTAAATCTTTTTCCTTCTGCTGCTTCATAGTGGCAGCTGTGTGTTAAAG 279
 QY 2024 caataatggtcagctcct 2083
 Db 338 CAATAATGTCAGTCCCTCATCTCTGCTGCTTCATAGTGGCAGCTGTGTGTTAAAG 279
 QY 2084 aaggaagcttgctacctttaaagtgtagtggaataaataaataaataaataaataaataa 2143
 Db 278 AAGGGAAGCTTTGACCTTTAGAGTGTAGGTGAATGAATGAATGAATGAATGAATGAATGA 219
 QY 2144 agaacagcatatgattcttcttcttcttcttcttcttcttcttcttcttcttcttcttct 2203
 Db 218 AGAACAGCATATGATTTCTTCTGCTTTGGGGAAAAAAGAAATGATGCTATGAAATGGTGGTG 159
 QY 2204 gtattttgagaagaataacattgtcttctctctctctctctctctctctctctctctctct 2263
 Db 158 GTGATTTGAGAACATTAATCATTTGCTTATGTCATTAATGAGCTGAATTTGATAAAACCCA 99
 QY 2264 aatacagcatatgaagtgtctgggcaagtttactctttttctctctctctctctctctct 2320
 Db 98 AAATACAGCTATGAAGTGTGGGCAAGTTTACTTTTTTTCTGATGTTTCTCTACAACT 42

RESULT 9

LOCUS BG289777
 DEFINITION 602385071F1 NIH_MGC_93 Homo sapiens cDNA clone IMAGE:4514129 5',
 mRNA sequence.
 ACCESSION BG289777
 VERSION BG289777.1 GI:13045907
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1. (bases 1 to 881)
 NIH-MGC <http://mgi.nci.nih.gov/>.
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 CONTACT: Robert Strausberg, Ph.D.
 Email: cagaps-remail.nih.gov
 Tissue Procurement: ATCC

found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: L1AM10407 row: m column: 12
 High quality sequence stop: 746.

FEATURES

source
 1. .882
 Location/Qualifiers
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone_lib="NIH_MGC_93"
 /tissue_type="transitional cell papilloma, cell line"
 /lab_host="DH10B (phage-resistant)"
 /note="Organ: bladder; Vector: pCMV-SPORT6; Site_1: NotI;
 site_2: SalI; Cloned unidirectionally; oligo-dT primed.
 Average insert size 1.7 kb. Library enriched for
 full-length clones and constructed by Life Technologies.
 Note: this is a NIH_MGC Library."
 BASE COUNT 259 a 155 c 202 g 266 t
 ORIGIN

Query Match 31.1%; Score 722.4; DB 10; Length 882;
 Best Local Similarity 94.5%; Pred. No. 3.3e-178;
 Matches 792; Conservative 0; Mismatches 41; Indels 5; Gaps 4;
 Qy 1447 gacctcaactggtacacagtcacattccttggagcgtgattggttctcttggccgtcgtg 1506
 Db 1 GACCTCACCTGGTACCAGTACCATTCCCTTGGAGCGTATTGGTTCTCTTGGCCGCTCGTG 60
 Qy 1507 ctgacagtgccctcatcacctttaaattgttgcttggcttggctacccggaatcttgggg 1566
 Db 61 CTGACAGTGGCCTTCATCACCCTTAAATGTTGTGCTTATGGTACCGGAATGCTTGGGG 120
 Qy 1567 aaaaaggcgagtgaaagaaacccacaaatccaaagaccattgagaagtggtgggaaa 1626
 Db 121 AAAAAAGGCGAGTTAAGAAAGCCCAAAATCCAAAGACCATTTGAGAAGTGGTGGGAAA 180
 Qy 1627 taaggtaaaatttggaaccattccctagtcatttccaaacttgaaacagaaatcagtgtt 1686
 Db 181 TAAGGTAAATTTTGAACCATTCCTCTAGCTATTTCCAAACTTGAACACAGATCACTGTT 240
 Qy 1687 aaattcatttattcttatttaaggaaatactttgcatgataatttaacagccccagagtgct 1746
 Db 241 AAATTTCATTTTATTTATTAAGGAAATACTTTTGCATAAATTAATTCAGCCCCAGAGTGCT 300
 Qy 1747 ttaaaaattctcttaataaaataatagactcgctagtcagtaaaagatatitgaaat 1806
 Db 301 TTAATAAATTCCTTAAATAAATAATAGACTCGCTAGTCACTAGTAAGATATTTGAATAT 360
 Qy 1807 gtatcgccccctccggtgtctttgatcagatgacatgtgccatttttcaagagacgt 1866
 Db 361 GTATCGTCCCCCTCCGGTGTCTTTGATCAGGATGACATGTGCCATTTTTCAGAGACGT 420
 Qy 1867 gcagacagctggcattctagattacttttcttactctgaacatggccctggttggagt 1926
 Db 421 GCAGACAGCTGGCATTTCTAGATTACTTTTCTACTCTGAACATGGCTGTTGGAGT 480
 Qy 1927 ggggattcaagggtggccacccgctgcccactactgcaaatggcaggttttaatatctc 1986
 Db 481 GCGGATTCAAAGGTGGTCCACCGCTGCCCTACTGCCAATGGCAGTTTAACTTATTC 540
 Qy 1987 ttttgctctgcagatgggtgcaattgatcttcaacaaatagttcagtcctcatctc 2046
 Db 541 TTTTGGCTCTGCAGATGGTTCGATTTGATTCCTTAACCAATATGGTCACTCTCATCTC 600
 Qy 2047 tgcctgttccatagggtccaccttgggt-gtttaagaagggaagcttggaccctttag 2105
 Db 601 TGTCCTGCTTCATAGTGGCCACCTTGTGTGGTTTAAAGAAAGGAAAGCTTTTGACCTTTAG 660
 Qy 2106 agtgtagtg-gaataagaatgaatggttggagtgacatgagaaacagcatatgattct-t 2163
 Db 661 AGTGTAGGTGAATGAATGTTGGAGGTGCACCTGAGAACATGATATGATTCTCTCG 720

Qy 2164 gctttgggaaaaaagaatgatgtatgaattggtggtggtgtatttgagaagataatc 2223
 Db 721 GCTTTGGGAAAAAAGAAATGATGCTATGCAAAATGGGGTGGGATTTGGAGAAAAATATC 780
 Qy 2224 attgcttatgtcacaatggagctgaatttgataaaaccccaaaatacacagctatgaagt 2281
 Db 781 ATGCTTTATGTCAAGAGAGG--TGATTGTATAAAACCAATACAGGTATTAAAGGTG 836
 RESULT 6
 BI759297
 LOCUS 603042939F1 NTH_MGC_116 Homo sapiens cDNA clone IMAGE:5183433 5',
 DEFINITION mRNA sequence.
 ACCESSION BI759297
 VERSION BI759297.1 GI:15750875
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 876)
 AUTHORS NIH-MGC http://mgc.nci.nih.gov/
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-r@mail.nih.gov
 Tissue Procurement: Life Technologies, Inc.
 CDNA Library Preparation: Life Technologies, Inc.
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: L1AM11457 row: n column: 10
 High quality sequence stop: 735.

FEATURES

Location/Qualifiers
 1. 876
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:5183433"
 /clone_lib="NIH_MGC_116"
 /lab_host="DH10B"
 /note="Organ: pooled colon, kidney, stomach; Vector:
 pCMV-SPORT6; Site_1: NotI; Site_2: EcoRV (destroyed); RNA
 source anonymous pool of 3 colons, age 26 yo male, 49 yo
 female, 71 yo male colon; 46 yo male kidney, and pool of 2
 stomachs, 62 yo male and 70 yo female. Library is
 oligo-dT primed and directionally cloned (EcoRV site is
 destroyed upon cloning). Average insert size 1.4 kb,
 insert size range 1-3 kb. Library is normalized and
 enriched for full-length clones and was constructed by C.
 Gruber (Invitrogen). Research Genetics tracking code
 023. Note: this is a NIH_MGC Library."
 BASE COUNT 226 a 193 c 205 g 252 t
 ORIGIN

Query Match 31.1%; Score 720.8; DB 10; Length 876;
 Best Local Similarity 94.6%; Pred. No. 8.6e-178;
 Matches 823; Conservative 0; Mismatches 37; Indels 10; Gaps 7;

Qy 1 ctcaactgcagttctctgatgcttgacaggtggacagcccccctctctatgttg 60
 Db 7 CTCAGCTCGAGTTCCTGTGAGGCTTGCACAGGTGGACGCCCCCTTCCTCTATGTGG 66
 Qy 61 tgcctgtcgtgcacctgtggtttggcagaggggaagactgagtgcccatggat 120
 Db 67 TGTCTGCTGTGACCTGTGGCTTTGCCAGCAGGCAAGCTACTGGTAGTGCCATGGAT 126
 Qy 121 gggagccactgggttccaccatgagtcggtgtggaagaaactcattctcagggggcagaag 180
 Db 127 GGGAGCCCACTGGTTCCACCATGATGAGTGGTGGTGGAGAAACTCATTTCTCAGGGGCAATGAG 186

source

```

1..2450
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:3854370"
/tissue_type="Colon, adenocarcinoma"
/clone_lib="NIH_MGC_65"
/lab_host="DH10B"
/notes="vector: pCMV-Sport6"
669 a 509 c 565 q 707 t
BASE COUNT

```

Query Match 70.3%; Score 1631.6; DB 11; Length 2450;
Best Local Similarity 82.6%; Pred. NO. 0;
Matches 1880; Conservative 0; Mismatches 394; Indels 2; Gaps 1;

Qy	47	ttcctctatgtgtgtgtctgtctgtacccctgtgctttgccgagcgagcgaggaagactactcg	106
Db	128	TTTTCTGACGGGGTTTCTTTTAGCACTTTGGGCAATGGTGTAGGTGACAAAGCTGCTG	187
Qy	107	tagtgcccatggatggagaccacttggttcaccatgagtcggtgggagaaaactcattc	166
Db	188	TGGTCCCTCAGGACGGAAGCCACTGGCTAGTATGAAGGATATAGTTGAGGTTCTCAGTG	247
Qy	167	tcagggggcataggtggttagtcatcccaagaggtgaattggcaactgggaagatcac	226
Db	248	ACCGGGGTCAATGAGATTGTAGTGGTGCGCTGGAAGTTAAATTGCTTTTGAAGAAATCCA	307
Qy	227	tgaattgcacagtgaagacttatccaacttcataccctcggaggtctggaccggaggt	286
Db	308	AATACTACACAAGAAAATCTATCCAGTCCGTATGACCAAGAAGAGCTGAAGAACCGTT	367
Qy	287	tcagggctttgccatgcctaattgg-aaagacaagtcgaagtatatattctctatt	344
Db	368	ACCAATCATTTGAAGAAACAACTCATCTGAGCGCATTCCTTAACCTCTCTCAGACAGA	427
Qy	345	aatgggttcatacaatgcatttttgactattttttcacaattgcagagtttgttttaa	404
Db	428	GTACAGAAATAACATGATGTGTTATTTGGCCTGTACTTCATCAACTGCCAGAGCCTCTGCA	487
Qy	405	agacaaaaaatagtagaataacttaaagagaggtcttttgatgcaggtgttctcgatcc	464
Db	488	GGACAGGGACACCCGTCACTTCCTTTAAGGAGAGCAAGTTTGATGCGTCCTTTTCACAGACC	547
Qy	465	ttttgataactgtggtgttaattgttgccaaaatattcttcctccctcccggtgtcttcgc	524
Db	548	AGCCTTACCCTGTGGGGTGATCCTGGGTGAGTATTTGGGGCTACCACTGTGTACCTCTT	607
Qy	525	caggggaataactttggcactatttgaaagaaggtgcacagtgccctgtctctcttcta	584
Db	608	CAGGGGTTTCCTGTGCTCCCTGGAGCATACATTCAGCAGAAGCCAGACCTGTGTCTTA	667
Qy	585	tgtcccaagaattctcttaggtgttctcagatgccactgacttcaaggagagagtaacgaa	644
Db	668	CATTCCCAGGTGTACACAAAGTTTTCAGACACATGACTATTTTCCCAACAGAGTGGCAA	727
Qy	645	ccacatcatgcacttgaggaacaattattatgcccacggtttttcaaaaatgcccctaga	704
Db	728	CTTCTTGTGTTAATTGTTGGAGCCCTATCTATTATTTATTTGTCGTTTTCAAGTATGAAGA	787
Qy	705	aatagcctctgaaattctctcaaacacccctgttacggagtatgtctctacagccacatc	764
Db	788	ACTCGCATCAGCTGCTCTCAAGAGATGTGGATAATAATCACCTTATATCAGAAGGTCTC	847
Qy	765	aatttggttgtgcgaagcgacttgtttgactatcccaaaacccgtgatgcccacat	824
Db	848	TGTTTGGCTGTTAAGATATGACTTTTGTGCTTGTAATATCCTAGGCCGGTCTGCCACAT	907
Qy	825	gacttcaattggtgttatcaactgccatcaggggaagccgttgctcatgtgaatttgaagc	884
Db	908	GGTCTTCATTGGAGGTATCACTGTAAAGAAGAGGAAGACATTTGTCACAGAAATTGAAGC	967
Qy	885	ccacttaactgctcttggaagacgaattgtgtgttctctcttggagatcaattgcttc	944

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 21, 2002, 09:38:48 ; Search time 4821.79 Seconds
(without alignments)
6494.048 Million cell updates/sec

Title: US-09-835-082-3
Perfect score: 2320
Sequence: 1 ctccagctgcagttctctgat.....ttctgatgttctctcaact 2320

Scoring table: IDENTITY_NUC
Gap 10.0 , Gapext 1.0

Searched: 13736207 seqs, 6748477542 residues
Total number of hits satisfying chosen parameters: 27472414

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*
1: em_estba.*
2: em_esthum.*
3: em_estin.*
4: em_estmu.*
5: em_estov.*
6: em_estpl.*
7: em_estro.*
8: em_hic.*
9: gb_est1.*
10: gb_est2.*
11: gb_hic.*
12: gb_gss.*
13: em_gss_hum.*
14: em_gss_inv.*
15: em_gss_pln.*
16: em_gss_vrt.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1631.6	70.3	2450	11 BC011409	BC011409 Homo sapi
2	789.6	34.0	2306	11 BC012716	BC012716 Mus muscu
3	780.2	33.6	2202	11 AK002629	AK002629 Mus muscu
4	741.2	31.9	868	10 BG402908	BG402908 602418710
5	722.4	31.1	882	10 BG291462	BG291462 602387347
6	720.8	31.1	876	10 BI759297	BI759297 603042939
7	707.4	30.5	884	10 BI762347	BI762347 603049004
8	695.8	30.0	818	9 AI307269	AI307269 tb24d12.x
9	695.4	30.0	881	10 BG289777	BG289777 602385071
10	680.6	29.3	1005	10 BE870911	BE870911 601448862
11	679.2	29.3	697	10 BG747143	BG747143 602704364
12	669.4	28.3	764	9 AI831562	AI831562 wj39601.x
13	644	27.8	723	10 BG424452	BG424452 602447772
14	641.2	27.6	829	10 BG469506	BG469506 602532818
15	640	27.6	833	10 BG400591	BG400591 602464244
16	639.8	27.6	791	9 AI765716	AI765716 w183f06.x
17	636.8	27.4	658	9 AV652040	AV652040

18	635.4	27.4	945	10 BG167885	BG167885 602339857
19	625.6	27.0	724	10 BF031444	BF031444 601559244
20	621.8	26.8	788	10 BG286763	BG286763 602382611
21	618.8	26.7	796	10 BG432363	BG432363 602496688
22	604.8	26.1	861	9 AI829268	AI829268 wk58a10.x
23	601	25.9	647	9 AV632287	AV632287 AV692287
24	596	25.7	763	10 BG399019	BG399019 6024400878
25	594.2	25.6	663	9 AV646537	AV646537 AV646537
26	592.4	25.5	859	10 BE249993	BE249993 600943171
27	589.8	25.4	640	9 AI635995	AI635995 w14e12.x
28	588	25.3	683	9 AI480360	AI480360 tm51d11.x
29	584.8	25.2	812	10 BG401258	BG401258 602465359
30	584	25.2	621	9 AW151709	AW151709 xf58b10.x
31	583	25.1	667	9 AW051715	AW051715 wx27g08.x
32	583	25.1	969	10 BG291839	BG291839 602386068
33	572	24.7	636	9 AI767470	AI767470 wh2a05.x
34	567.4	24.5	581	10 BE379923	BE379923 601159412
35	567	24.4	623	10 BG171502	BG171502 602322042
36	557.8	24.0	743	10 BG386668	BG386668 602454928
37	556.4	24.0	741	10 BG433491	BG433491 602497921
38	550.6	23.7	619	9 AW300312	AW300312 xs9f05.x
39	549.8	23.7	851	10 BG401340	BG401340 602465455
40	543.8	23.4	700	10 BF130433	BF130433 601818721
41	542.8	23.4	608	9 AW469219	AW469219 hc79e05.x
42	542.8	23.4	953	10 BG328092	BG328092 602427143
43	537	23.1	935	10 BF234991	BF234991 602027544
44	535.4	23.1	626	9 AW082933	AW082933 xc04f05.x
45	532.2	22.9	666	10 BF027719	BF027719 601764020

ALIGNMENTS

RESULT	1
LOCUS	BC011409 2450 bp mRNA linear HTC 30-JUL-2001
DEFINITION	Homo sapiens, clone IMAGE:3854370, mRNA.
ACCESSION	BC011409
VERSION	BC011409.1 GI:15030273
KEYWORDS	HTC.
SOURCE	human.
ORGANISM	Homo sapiens
REFERENCE	1 (bases 1 to 2450)
AUTHORS	Strausberg, R.
TITLE	Direct Submission
JOURNAL	Submitted (25-JUL-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA
REMARK	NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT	Contact: MGC help desk Email: cgapbs-r@mail.nih.gov Tissue Procurement: ATCC cDNA Library Preparation: Life Technologies, Inc. cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA Sequencing by: Baylor College of Medicine Human Genome Sequencing Center Center code: BCM-HGSC Web site: http://www.hgsc.bcm.tmc.edu/cdna/ Contact: villalon@bcm.tmc.edu Villalon, D.K., Luna, R.A., Hale, S.M., Hulyk, S., Lu, X., Garcia, A.M., Holloway, M., Telford, B., Hodgson, A., Bouck, J., Yu, W., Muzny, D.M., Gibbs, R.A.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Series: IRAK Plate: 13 Row: b Column: 4
This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 4507814
This clone has the following problem: incomplete processing.
Location/Qualifiers


```

; APPLICANT: BELKNAP, WILLIAM R
; TITLE OF INVENTION: DNA SEQUENCES ENCODING SOLANIDINE
; TITLE OF INVENTION: UDP-GLUCOSE GLUCOSYLTRANSFERASE AND USE TO REDUCE
; TITLE OF INVENTION: GLYCOALKALOIDS IN SOLANACEOUS PLANTS
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NANCY J. PARSONS
; STREET: 800 BUCHANAN ST.
; CITY: ALBANY
; STATE: CA
; COUNTRY: USA
; ZIP: 94710
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/797,226
; FILING DATE:
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: PARSONS, NANCY J
; REGISTRATION NUMBER: 40,364
; REFERENCE/DOCKET NUMBER: 0011.97
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (510) 559-5731
; TELEFAX: (510) 559-5777
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1607 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA to mRNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Solanum tuberosum
; STRAIN: cv. Lemhi Russet
; INDIVIDUAL ISOLATE: SGT1750
; DEVELOPMENTAL STAGE: mature, somatic
; TISSUE TYPE: tuber
; IMMEDIATE SOURCE:
; LIBRARY: lambda gt11 cDNA library
; CLONE: SGT 1750
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 20..1486
; OTHER INFORMATION: /product="solanidine
; OTHER INFORMATION: glucosyltransferase"
; PUBLICATION INFORMATION:
; AUTHORS: Moehs, Charles P.
; AUTHORS: Allen, Paul V.
; AUTHORS: Friedman, Mendel
; AUTHORS: Belknap, William R.
; TITLE: Cloning and expression of solanidine
; TITLE: UDP-glucose glucosyltransferase from potato
; JOURNAL: The Plant Journal
; VOLUME: 11
; ISSUE: 2
; DATE: 1997
; RELEVANT RESIDUES IN SEQ ID NO: 1: FROM 1 TO 488
; US-08-797-226-1

```

```

Query Match 1.9%; Score 44.8; DB 2; Length 1607;
Best Local Similarity 53.4%; Pred. No. 0.0072;
Matches 94; Conservative 0; Mismatches 82; Indels 0; Gaps 0;

QY 1068 gtgggtaccccaaaacatctgctgtgaccccgatgacccgtgctcttatcacccatgc 1127
      ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1069 GTGGTCCACAGCTTACGATCATGGAACATTCAGCAACAGCGGGTTCATGACTCATG 1128

```

```

QY 1128 tggttcccatggtgtttatgaagcatatgcaatggcgttcccatggtgatgacccctt 1187
      ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1129 TGGTAACTAATTCGGTTCGGAAGCCATCACTTTTGGCGTGCATGATACATGCGCCACT 1188
      ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1188 gtttggtgacagatggaacaatgcaaaagcgcacatgagagactaaaggaagctggagtga 1243
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1189 TTATGCTGATCAATCTCAACAGAGAAGTAGTCGAGGTTAGGGGATTGGGAATCA 1244
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 15
PCT-US92-00282-37
; Sequence 37, Application PC/TUS9200282
; GENERAL INFORMATION:
; APPLICANT: OWENS, IDA S.
; APPLICANT: RITTER, JOSEPH K.
; TITLE OF INVENTION: THE GENETIC LOCUS UGT1 AND A MUTATION
; TITLE OF INVENTION: THEREIN.
; NUMBER OF SEQUENCES: 40
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CUSHMAN DARBY & CUSHMAN
; STREET: 1615 L STREET, N.W.
; CITY: WASHINGTON
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20036-5601
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US92/00282
; FILING DATE: 19920110
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: SCOTT, WATSON T.
; REGISTRATION NUMBER: 26581
; REFERENCE/DOCKET NUMBER: 91532-PCT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-861-3000
; TELEFAX: 202-822-0944
; TELEX: 6714527 CUSH
; INFORMATION FOR SEQ ID NO: 37:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 51 base pairs
; TYPE: NUCLEIC ACID
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; PCT-US92-00282-37

Query Match 1.9%; Score 43.6; DB 5; Length 51;
Best Local Similarity 92.0%; Pred. No. 0.0021;
Matches 46; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1444 cacagacctacactgtgtacacagatcaccattccttgagcgtgattggttccct 1493
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1 CACGACCTCACCTGGTCCATCCACTCTTTGGATGTGATGTTGTTCT 50

```

Search completed: August 21, 2002, 11:59:49
Job time: 32837 sec

;
; TYPE: NUCLEIC ACID
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cdna
PCT-US92-00282-18

Query Match 7.0%; Score 161.8; DB 5; Length 1190;
Best Local Similarity 51.8%; Pred. No. 1.3e-36;
Matches 414; Conservative 0; Mismatches 382; Indels 3; Gaps 2;

```
Qy 68 tctgacacgttggttgcgagcaggaagctactgtagtgcacatgagtgagacc 127
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 140 TGCTGGGCCCGCTGCTCCCATGCTGGGAAGATAGTGTGATCCAGTGGATGGCACC 199
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 128 actggttcaaccatgaagtcggtggtgagaaactcatttcagggggcagtgagtggtg 187
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 200 ACTGCTGAGCATGCTTGGGGCCATCCAGCAGCTGCAGCAGAGGGGACATGAATAGTTG 259
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 188 tagtcacgcagaggtgagttggcaactgggaagatcactgaattgcacagtgaaagactt 247
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 260 TCCTAGCACCTGAGCGCTGTTGTATCATCAGAGCGGACATTTTACACCTTGAAGCGT 319
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 248 atteaactcatatccctggagatctgacccggagtt--caagccttttgcccatgc 305
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 320 ACCCTGTGCCAATCCAAAGGGAGGATGTGAAGAGTCTTTTGTAGTCTCGGGCATAATG 379
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 306 tcaatggaagacaagtagtgaagtatatattctctatttaattgggttcatacaatgacat 365
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 380 TTTTGTGAGATGATCTTTCCTGCGAGCGTGTGATCAAAACATACAGAAATAAANAAGG 439
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 366 ttttgact-tattttttcaaatgacggaggtgtttaaagacaaataattagtagaat 424
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 440 ACTCTGTATGCTTTTGTCTGGCTGTTCACACTTACTGCAACAAGAGAGCTCATGGCT 499
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 425 acttaagagaggtctttttgatgcaggtttctctgatcctctttgataaactgtgcttaa 484
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 500 CCTGGCAGAAAGCAGCTTTGATGTCATGCTGAGCGAGCCCTTCTCTCTGCGAGCCCA 559
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 485 tgggtgcaaatcttctccctccctccctccctccctccctccctccctccctccctccct 544
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 560 TCGTGGCCAGTACCTGCTCTGCGCCACTGTATCTCTGATGCTGCTGCTGCTGCTGCTGCT 619
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 545 atcttgaaagagtcacagtcgctctctctctctctctctctctctctctctctctcttag 604
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 620 TGGAAATTGAGGCTACCCAGTGCCTCCCAACCCATCTCTGATGCTGCTGCTGCTGCTGCT 679
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 605 gttctcagatgcctgactttcaagagagagatcagaaacacacatcactcactggagg 664
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 680 CTCATTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 739
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 665 aacatttatgccacogtcttttttcaaaaatgcccctagaataagcctctgaaattctcc 724
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 740 AGAATCTTCTGTCGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 799
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 725 aacacactgttacgagtagtatctctacagccacacacatcattgtgttgcaagc 784
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 800 AGAGAGAGGTGATGCTCAGGACCTATTTAGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 859
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 785 actttgtttggactatccaaacccctgtagcccaacatgatcttctattgttggtatca 844
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 860 ACTTTGTGAAGGATGACCTAGGCCCATCATGCCCAATATGTTTGTGTTGTTGTTGTTGTT 919
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 845 actgcatcaggaaagacc 863
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 920 ACTGCTCTACCAAAATCC 938
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```

RESULT 10
PCT-US92-00282-12
; Sequence 12, Application PC/TUS9200282
; GENERAL INFORMATION:
; APPLICANT: OWENS, IDA S.

```
;  
; APPLICANT: RITTER, JOSEPH K.  
; TITLE OF INVENTION: THE GENETIC LOCUS UGT1 AND A MUTATION  
; THEREIN.  
; NUMBER OF SEQUENCES: 40  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: CUSHMAN DARBY & CUSHMAN  
; STREET: 1615 L STREET, N.W.  
; CITY: WASHINGTON  
; STATE: D.C.  
; COUNTRY: U.S.A.  
; ZIP: 20036-5601  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US92/00282  
; FILING DATE: 19920110  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: SCOTT, WATSON T.  
; REGISTRATION NUMBER: 26581  
; REFERENCE/DOCKET NUMBER: 91532-PCT  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 202-861-3000  
; TELEFAX: 202-822-0944  
; TELEX: 6714627 CUSH  
; INFORMATION FOR SEQ ID NO: 12:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1197 base pairs  
; TYPE: NUCLEIC ACID  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: cdna  
PCT-US92-00282-12
```

Query Match 6.9%; Score 161.2; DB 5; Length 1197;
Best Local Similarity 52.6%; Pred. No. 2e-36;
Matches 421; Conservative 0; Mismatches 373; Indels 6; Gaps 3;

```
Qy 85 gccgagcagggagactgtagtgcacatggatggagggagccactgg--ttcaaccatga 142
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 304 GCTGAGAGTGAAGAGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 363
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 143 ggtcgtggtggagaaactcattctcaggggcatgagtggtgttagtcagtcagcaag 202
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 364 GGGAGGCTTTCGGGAGTCTCCATGCCAGAGGCCACCGCGGTGGTCTTCCACCCAGAGG 423
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 203 tgagttggcaactgggaagatcactgaattgcacagtgaaagacttattcaacttcata 262
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 424 TGAATATGCACATCAAAAGAGAAATTTTTCACCTGCACAGCCTATGTGTTCATGGA 483
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 263 cccctggagatcgcacgggaggttaagcgttttgccatgctcaat--ggaaagcac 319
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 484 CCCAAGAGGAATTTGATCGCGTACGCTGGCTACACTCAAGGGTTCTTTGAAACAGAAC 543
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 320 aagtcagagatataattctctatttaattgggtttcacaatgacatttttgacttatttt 379
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 544 ATCTTCTGAAGAGATATCTTAGAAGTATGGCAATTTATGAACAATATCTTTTGGCCCTTC 603
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 380 tttaaatgacagaggttggtttaagacacaaaaaattagtagaataacttaagagagactt 439
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 604 ATAGTGTGTTGTGGAGCTACTGCAATATGAGGCCCTGATCAGGCACCTGATGATGATGAT 663
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 440 ettttgatgcagtgtttctctgatcctttttgataaactgtggcttaattgttgccaaatatt 499
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 664 CCTTTGATGTGTTTAAACAGACCCCGTTAACTCTGTGGGGCGGTGCTGCTGAAGTACC 723
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 500 tctccctccctccctccctccctccctccctccctccctccctccctccctccctccctcc 559
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 724 TCTCGATCTCTGCTGTGTTTGTGAGAGGTACATTCATGCTAGACTTATAGGCCA 783
```



```
Db 303 GGTGAATATGTACATCAAGAGAGAACTTTTTCACCCCTGACAAAGCTATGCCATTTCATG 362
Qy 261 taccctgagagatcgaccggaggttcaaggcttttgcocattgtcgaat---ggaagc 317
Db 363 GACCCAGGACGAATTTGATGGCTTTTCTGGGTCACTCAATCGTTCTTTGAAACAGA 422
Qy 318 acaagtagagatatactctctattatgggttcataaatgacattttgacttatt 377
Db 423 ACATCTCTGTATGAATTTCTAGAGAAATGCGCAATATGACAAATATGCTTTGTATCAT 482
Qy 378 ttttcaaatgagaggtttgtttaaagacaaataatagtagaataacttaagagag 437
Db 483 ACATAGGTCTGTGGAGCTACTGCATAATGAGGCCCTGATCAGGACCTGCTATAC 542
Qy 438 tcttttgcagctgtttctcgatcttttgaataactgtgcttaattgttccaaata 497
Db 543 TTCCTTGTATGTTCTTAACAGACCCCTTTTACCTCTGCGGGCGGTGCTGGCTAAGTA 602
Qy 498 tttctccctccctccgtgtgtcttcgaggggaataactttgccactatcttgaagaag 557
Db 603 CTGTGCGATTCTCTGTGTTTCTTGAGGAACATTCATGATTTAGACTTTTAAGGG 662
Qy 558 tgcagtgccctcctctctctctatgtcccaagaattctcttagggttctcagatgc 617
Db 663 CACACAGTGTCAAAACCTTCTCTCTATATCTCTAGATTAACAGCACCATTACAGCCA 722
Qy 618 catgactttcaagagagtagacggaacacacatcgtcacttggaggaaacattattatg 677
Db 723 CATGACATCTCTGAAAGGCTCAAGAACATGCTCTACCCCTGCGCCCTGTCTACCTTG 782
Qy 678 ccacgtttttcaaaatgccctagaaatagctctgaaattctccaaacacctttac 737
Db 783 CATGCTGTTCTCTCTTATGCAAGCTTGCCTCTGAGCTTTTCAGAGAGAGGTGC 842
Qy 738 ggaagtatgactctacagcacacacatcaatttgggttggcgaagcatttttttga 797
Db 843 AGTGTGTGATCTTCAGCAGCATGCTGCTGGTCTGTTCCGAGGGAGGTGTGTGATGA 902
Qy 798 ctatcccaaacccgtagtcccaacatgattctcattgtgtgtatcactgacctcaggg 857
Db 903 TTACCCAGGCGGATCATGCCAACATGGTCTTCATTTGGGGGCACTCACTGTGCCAACG 962
Qy 858 aaagcgttgctatggaattgaagcctacatta 892
Db 963 GAAGCCACTATCTCAGGTCTGTATTGTTGGTGCCTTTA 997
```

RESULT 7

```
; Sequence 15, Application PC/TUS9200282
; GENERAL INFORMATION:
; APPLICANT: OWENS, IDA S.
; TITLE OF INVENTION: RITTER, JOSEPH K.
; TITLE OF INVENTION: THE GENETIC LOCUS UGT1 AND A MUTATION
; NUMBER OF SEQUENCES: 40
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CUSHMAN DARB Y & CUSHMAN
; STREET: 1615 L STREET, N.W.
; CITY: WASHINGTON
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20036-5601
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US92/00282
; FILING DATE: 19920110
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
```

```
; NAME: SCOTT, WATSON T.
; REGISTRATION NUMBER: 26581
; REFERENCE/DOCKET NUMBER: 91532-PCT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-861-3000
; TELEFAX: 202-822-0944
; TELEX: 6714627 CUSH
; INFORMATION FOR SEQ ID NO: 16:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1448 base pairs
; TYPE: NUCLEIC ACID
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: YES
; PCT-US92-00282-16
```

```
Query Match 8.0%; Score 186.6; DB 5; Length 1448;
Best Local Similarity 53.8%; Pred. No. 1.1e-43;
Matches 429; Conservative 0; Mismatches 364; Indels 4; Gaps 2;
```

```
Qy 81 ctttcgagggcaggaagactactgttagtgcocattgagggagcactggttcaact 140
Db 178 CTGGGCTCTAGAGAGTGGGAAGGTGCTGGTGTACCCATTGATGGAGCCACCTGGGTG-CAT 236
Qy 141 gagggtcggtggtgagaaactcattctcaggggcatgaggtggtgttagtcatgcaga 200
Db 237 GCGGAGGCGCTGCGGGAGCTCCATGCCAGAGGCCACAGGGGGTGTCTCACCACCAGA 296
Qy 201 gttgagttggaactgggaagatcaactgaattgcacagtgaaagacttatcaactcata 260
Db 297 GGTGAATATGCATCAACAAAGAGAAATTTTCACCCCTGACACACCTTATGCCATTTCATG 356
Qy 261 taccctgagggatctggaccgggaggttcaaggcttttggcccatgctcaat---ggaagc 317
Db 357 GACCAGGATGAATTTGATCGCTTGTCTGGGCACACTCACTCACTGATCTTGTGAACAGA 416
Qy 318 acaagtacgaagtataattctctataatgggttccatacaatacaattttgacttatt 377
Db 417 ACATTTCTGACACATTTTCTAAAAGTATGGCAATTTTGAANAATTCATCTTTGGTCTT 476
Qy 378 ttttcaaatgagaggtttgtttaaagacaaataatagtagaataacttaagagagag 437
Db 477 TCATAGGTCTTGTGTGGAGCTACTGCATAATGAGGCCCTGATCAGGACACCTGAATGCTAC 536
Qy 438 ttttttgatcgaggtttctcgatccttttgataactgtgcttaattgttgcgaata 497
Db 537 TTCTTTCGATGGTGGTTTAAATAGACCCCAATTTACCTCTGTGGGCGAGTGTGCTAAGTA 596
Qy 498 tttctccctccctccgtgtgttcgocaggggaaataactttgcccactatcttgaagaag 557
Db 597 CCTATCAATTCCTGCTGTGTTTTTTTGGAGAACATTCATTCATGTGATTAGACTTTAAGGG 656
Qy 558 tgcacgtgcccgtcctctctctatgtcccaagaattctcttagggttctcagatgc 617
Db 657 CACACAGTGTCCAAATCCCTCTCTATATTCCTAAGTTACTTAAGCACAATTCACACCA 716
Qy 618 catgactttcaagagagtagtacgggaacacacatcattgacttggaggaacattattatg 677
Db 717 CATGACATTCCTGCAAAAGGGTCAAGAACATGCTCTACCCCTGTCCTCTACCTTTG 776
Qy 678 ccacgttttttcaaaaatgccctagaaatagctctgaaattctccaaacacctgttac 737
Db 777 CCATGCTCTTCTGCTCTTATGCAAGCCTTGCTCTGAGCTTTTTCAGAGGGAGGTGC 836
Qy 738 ggaagtatgactctacagccacacataatttgggttggtaacagcactttgttttga 797
Db 837 AGTGTGGATCTTCTCAGCCATGTCATGTGTGGCTGTTCCGAGGGGCACTTTGTGATGA 896
Qy 798 ctatcccaaacccgtagtcccaacatgattctcattgtgtgtatcaactgacctcaggg 857
Db 897 TTACCCAGGCGGATCATGCCAACATGGTCTTCATTTGGGGGCACTCACTGTGCCAACAG 956
```



```

QY 1505 tgcgtacagtgccctccatcccttaaatgtgtgtctatggtaccggaatcttgg 1564
Db 1496 TGCTGACAGTGGCTTCTATCATACCTTTAAATGTTGTCTATGGCTACCGGAATGCTTGG 1555
QY 1565 ggaauaaggcagtgtaagaaacccacaaatccaagaccattgagaagtgggtgga 1624
Db 1556 GGAANAAGGCGGAGTGAAGAAGCCCAAAATCCAAAGCCATTTGAGAAGTGGGTGGGA 1615
QY 1625 aataagtgaaattttgaaccattccctagctattcccaacttgaaacagaatcagt 1684
Db 1616 AATAGGTAAATTTTGAACCATTCCTAGTCAATTTCCAAACTTGAAGAACAGAAATCAGT 1675
QY 1685 ttaatacatttattcttataaggaatactttgcatataatcaatcagcccccagagt 1744
Db 1676 TTAATTCATTTTATCTTATTAAGGAATACCTTTGCAATAATTAATCAGCCCCAGAGTG 1735
QY 1745 ctttaaaaaattctcttaataaaaaataatagactcgtgtagtcagtaagaatattgaat 1804
Db 1736 CTTTAAAAAATTCCTTAAATAAAATAATAGACTCGCTAGTCAAGTAAGATATTGTAAT 1795
QY 1805 atgtatcgtgccctcccggtgctttgtacagagatgacatgtgccatttttcagagac 1864
Db 1796 ATGATNTGTCGCCCTCCGCTGCTTTGATCAGGATGACATGTCCTATTTTCAGAGAC 1855
QY 1865 gtgcagacaggtggtggaattctctagattacttttactctgaacatggtcgtgttggga 1924
Db 1856 GTGCAGACAGGCTGGCATTTCTAGATTACTTTTCTTACTCTGAAACATGGCTGTTTGGGA 1915
QY 1925 gtgcggattcaaggtggtccacgctgcccactgacacgaatgagtgagtttaattccta 1984
Db 1916 GTGCGGATTCAAAGTGGTCCACCGCTGCCCTACTGCAAAATGGCAGTTTAAATCTTA 1975
QY 1985 tcttttggcttcagatggttgcaattgactcttaaccataatggtcagtcctcatc 2044
Db 1976 TCTTGGCTTCTGCAAGATGGTGTCAATTTGATCCTTAACCAATAATGGTCAGTCTCATC 2035
QY 2045 tctgtcctgcttcataggtgcccacttgtgttttaaaagaggaagcttctgaccttta 2104
Db 2036 TCTGCTCTCTCATAGGTGCCACCTTGTGTGTTTAAAGAGGAGGCTTTGTACCTTTA 2095
QY 2105 aggtgtaggtgaaatgaatgaatggtgtgagtgacgtgacagacagatatgattcttg 2164
Db 2096 GAGTGTAGGTGAAATGAATGAGTGGCTGGAGTGCACGTGAGAACAGCATATGATTTCTTG 2155
QY 2165 ctttggggaagaaatgatgctatgaattggtggtggtgatttggagaagataatca 2224
Db 2156 CTTTGGGGAAGAAATGATGCTATGAATTTGGTGGTGGTGTATTTGAGAAGATAATCA 2215
QY 2225 tgccttatgcaatggagctgaatttgataaaaccccaaaatacagctatgaagtgtg 2284
Db 2216 TTGCTTATGTCAAATGGAGCTGAATTTGATATAAACCCAAATACAGCTATGAAGTGCTG 2275
QY 2285 ggcgaugttactttttcttctgatgttctcctcaact 2320
Db 2276 GCCAAGTTTACTTTTTTCTGATGTTTCTTCAAACT 2311

```

RESULT 3

PCR-US92-00282-25

Sequence 25, Application PC/US9200282

GENERAL INFORMATION:

APPLICANT: OWENS, IDA S.

APPLICANT: RITTER, JOSEPH K.

TITLE OF INVENTION: THE GENETIC LOCUS UGT1 AND A MUTATION

TITLE OF INVENTION: THEREIN.

NUMBER OF SEQUENCES: 40

CORRESPONDENCE ADDRESS:

ADDRESSEE: CUSHMAN DARBY & CUSHMAN

STREET: 1615 L STREET, N.W.

CITY: WASHINGTON

STATE: D.C.

COUNTRY: U.S.A.

```

; ZIP: 20036-5601
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US92/00282
; FILING DATE: 19920110
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: SCOTT, WATSON T.
; REGISTRATION NUMBER: 26581
; REFERENCE/DOCKET NUMBER: 91532-PCT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-861-3000
; TELEFAX: 202-822-0944
; INFORMATION FOR SEQ ID NO: 25:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1561 base pairs
; TYPE: NUCLEIC ACID
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; PCT-US92-00282-25

```

Query Match 41.7%; Score 967.2; DB 5; Length 1561;

Best Local Similarity 98.0%; Pred. No. 3.3e-268;

Matches 1011; Conservative 0; Mismatches 18; Indels 3; Gaps 3;

```

QY 1291 aaagcagctcatcaatgacaaagttacaaggagaacatcatgagcctctccagcctccac 1350
Db 248 AACGAAACTTCTCTTTGTTGTTTACAGAGAAACATCATGCGCTCTCCAGCCTTCAC 307
QY 1351 aaggaaccgcscggtgagcgctgacctggcgctgtctctgttctgggtggagttgtgatgag 1410
Db 308 AAGGACCGC:CGGTGGAGCGCGTGGACCTGGCGGTCTCTGGGTGGAGTTGTGTGATGAGG 367
QY 1411 cacaagggcgccacacactcgcccgagccccagacctcaacctggtaccagtaaccat 1470
Db 368 CACAAGGGCGCGCACACCTCGGCCCGCGAGCCACGACCTCACCTGGTACCAGTACCATT 427
QY 1471 tccctggacgtgattggttctctctcttggcgctgtgct-gacagtggtccttcacacct 1529
Db 428 TCCCTGGACGTGATTGGTTCTCTCTTGGCGGTCTGGCTGGACAGTGGCCCTTCATCACCTT 487
QY 1530 taaatgttgccttatgctaccggaataatgcttgggaaaaaagggcgaggttaagaaagc 1589
Db 488 TAAATGTTGTCTTATGGCTACCGAAATGCTTGGGAAAAAAGGGCGAGTTTAAAGAACG 547
QY 1590 ccacaaatccaagaccattgagaagtgggtgggaaataaggtaaaaattttgaaccattc 1649
Db 548 CCACAAATC:AAAGCCCATTTGAGAAAGTGGTGGGAAATAAGGTAAAAATTTTGAACCATTC 607
QY 1650 cctagtcatcttcaaacctgaaacaaatcagtgtaaatcttatttcttcttcttaag 1709
Db 608 CCTAGTCTTTCCAACTTGAAACAGAAATCAGTGTAAATTCATTTTATCTTATTTAG 667
QY 1710 gaaatactttgcataaaataatcagccccagagtgcttttaaaaaattctcttaataaaa 1769
Db 668 GAAATACTTTGCAATAAATAATCAAGCCCGCAGAGTGTCTTAAAAAATTTCTCTTAAATAAAA 727
QY 1770 ataataactcgcctagtcagtaaaagatatattgaatatatcgtg-ccccctccggtgtc 1828
Db 728 ATAATAGACTCGCTAGTCAGTAAAGATATTTGAAATATGATGTCGCCCCCTCTGCTGTGTC 787
QY 1829 tttgatcagatgacatgtgccatttttccagagagcgtgcagacagcgtggcattctaga 1888
Db 788 TTTGATCAG:ATGACATGTGCCATTTTTCAGAGGACGTGCAGACAGAGCTGG-ATTCTAGA 846
QY 1889 ttacttttctactctgaacacatggtcctgtttgggagtgcggttcaaaagtggtccca 1948

```


QY	1252	gtctggaatgactcttgaagatttagaataagtctataaagcagtcacatgaacaa	1311
Db	421	gtctctgaaatgactgtgatgttggaaatgccttaaactgtcatcaacaag	480
QY	1312	agttacaaggagacaatcatcgctctccagcttccaaaggacgcgcggtagccg	1371
Db	481	agttacaaggagacaatcatcgctctccagcttccaaaggacgcgcgtctatagcct	540
QY	1372	ctggacctggccgtgttcttgggtggagtgtgtgatgagcacaaggcgcgcacactg	1431
Db	541	ctggacctggcgtgttcttgggtggagtgtgtgatgagcacaaggcgcacacactg	600
QY	1432	cgcccgacgcccacgacctccactgtgtacacatgccattccttggagtgatgtgttc	1491
Db	601	cgcccgcgccacctgacctccactgtgtacacatgccattccttggatgtgtgtgttc	660
QY	1492	ctctctggccgcgtgtctgacagtgcctctcatcacctttaatgtgtcttatggctac	1551
Db	661	ctctctggcccatgtgttgacagtggtcttcatctgacctttaaattgtgtccatggctgc	720
QY	1552	cggaaatgcttggggaaaaaaggcgaggttaagaaaggccacaactccaagaccattga	1611
Db	721	cggaaatgcttgggggaaaggcgcnagtgaagaaatcacacaatccaagaccattga	780
QY	1612	gaagtgggtgggaaataaagtgaaaaatttgaacattccct	1652
Db	781	gaagtgggggaagtgaangagaagatattagtttcattct	821

Search completed: August 21, 2002, 12:46:07
Job time: 22285 sec

CC animals or modified cells e.g. for pharmacogenetic screening.

XX Sequence 759 BP; 190 A; 165 C; 169 G; 235 T; 0 other;

Query Match 28.4%; Score 659.8; DB 21; Length 759;
Best Local Similarity 91.8%; Pred. No. 4.4e-177;
Matches 697; Conservative 0; Mismatches 62; Indels 0; Gaps 0;

Qy 115 atggatggagccactgttcacacatgagtcggtggtgagagaaactcattcagggg 174
Db 1 atggatggagtcactgttcacacatgagtcggtggtgagagaaactcattcagggg 60
Qy 175 catgagggtgtgtgacatgacagatgagtcggtgagagaaactcattcagggg 234
Db 61 catgagggtgtgtgacatgacagatgagtcggtgagagaaactcattcagggg 120
Qy 235 acagtgaagacttattacacttcacacatgagtcggtgagagaaactcattcagggg 294
Db 121 acagtgaagacttattacacttcacacatgagtcggtgagagaaactcattcagggg 180
Qy 295 ttgcccactgtcactgaagacacacatgagtcggtgagagaaactcattcagggg 354
Db 181 ttgcccactgtcactgaagacacacatgagtcggtgagagaaactcattcagggg 240
Qy 355 tacaatgacattttgacttatttttccaaatgagtcggtgagagaaactcattcagggg 414
Db 241 tccagtggtttctgacttatttttccaaatgagtcggtgagagaaactcattcagggg 300
Qy 415 ttgtagaataacttaagagagatcttttgcagtcggtgagagaaactcattcagggg 474
Db 301 ttgtagaataacttaagagagatcttttgcagtcggtgagagaaactcattcagggg 360
Qy 475 tgggttaattttgacttatttttccaaatgagtcggtgagagaaactcattcagggg 534
Db 361 tgggttaattttgacttatttttccaaatgagtcggtgagagaaactcattcagggg 420
Qy 535 ctttgcactatctgaagagatgacagtcggtgagagaaactcattcagggg 594
Db 421 ttttgcactatctgaagagatgacagtcggtgagagaaactcattcagggg 480
Qy 595 attctctaggggttcagatgacatgacttccaaatgagtcggtgagagaaactcattcagggg 654
Db 481 gatctcttaggggttcagatgacatgacttccaaatgagtcggtgagagaaactcattcagggg 540
Qy 655 cacttgaggagacatttatttgcacacggttttttccaaatgagtcggtgagagaaactcattcagggg 714
Db 541 cacttgaggagacatttatttgcacacggttttttccaaatgagtcggtgagagaaactcattcagggg 600
Qy 715 gaaatttccaaacactgttaccgagtcggtgagagaaactcattcagggg 774
Db 601 gaaatttccaaacactgttaccgagtcggtgagagaaactcattcagggg 660
Qy 775 ttgcgaacgactttgtttgactatcccaacccgtgagtcggtgagagaaactcattcagggg 834
Db 661 ttgcgaacgactttgtttgactatcccaacccgtgagtcggtgagagaaactcattcagggg 720
Qy 835 ggtgtatcaactgacacagggagacggtgagtcggtgagagaaactcattcagggg 873
Db 721 ggtgtatcaactgacacagggagacggtgagtcggtgagagaaactcattcagggg 759

RESULT 15
ABL01468
ID ABL01468 standard; DNA; 874 BP.
XX
AC ABL01468;
XX
DT 15-MAR-2002 (first entry)
XX Murine apoptosis related DNA sequence #133.
DE
XX Apoptosis; mouse; cancer; autoimmune disease; viral infection;
KW

KW Alzheimer's disease; Parkinson's disease; Huntington's disease;
KW reperfusion injury; stroke; liver damage; dilatory cardiomyopathy;
KW transgenic animal; hepatotropic; antialcoholism; cytostatic;
KW immunosuppressive; virucide; neurotropic; neuroprotective; vasotropic;
KW antiparkinsonian; cerebroprotective; ds.
XX Mus sp.
OS
PN DE10126344-AL.
XX
XX 24-JAN-2002.
XX
XX 30-MAY-2001; 2001DE-1026344.
XX
XX 14-JUL-2000; 2000DE-1034303.
XX
XX (PLAC) MAX PLANCK GBS FOERDERUNG WISSENSCHAFTEN.
XX
XX Grimm S, Schoenfeld N, Braziulis E, Cramer U, Gewies A, Voss F;
PI Mund T, Albayrak T, Gille H, Klein M;
XX
XX WPI; 2002-115563/16.
XX
XX New apoptosis-associated nucleic acid sequences and polypeptides,
XX useful for diagnosis, treatment and prevention of e.g. tumors and
XX neurodegeneration -
XX
XX Claim 1; Page 184; 227pp; German.

XX The present invention relates to nucleic acids from the mouse, where the
XX nucleic acid is associated with apoptosis. The sequences can be used in
XX the diagnosis, treatment and prevention of diseases associated with
XX excessive or inadequate apoptosis, including tumours, autoimmune
XX diseases, viral infections, degenerative diseases (Alzheimer's,
XX Parkinson's and Huntington's diseases), reperfusion injury, stroke and
XX alcohol-induced injury to the liver, for identifying agents for treating
XX these diseases, and to prepare transgenic animals in which expression of
XX an apoptosis related sequence is altered. These are useful for genetic
XX and/or pharmacological investigations of apoptosis and related diseases,
XX including dilatory cardiomyopathy. The present sequence is one of the
XX apoptosis related sequences of the invention.

XX Sequence 874 BP; 222 A; 202 C; 233 G; 210 T; 7 other;

Query Match 24.9%; Score 578.8; DB 24; Length 874;
Best Local Similarity 84.9%; Pred. No. 5.3e-154;
Matches 646; Conservative 0; Mismatches 115; Indels 0; Gaps 0;

Qy 892 aatgctctggagaaacatggaattgtggtttctcttgggataatggtctcagaatt 951
Db 61 aacgctctggagaaacatggaattgtggtttctcttgggataatggtctcagaatt 120
Qy 952 ccagagaagaacatggaattgtggtttctcttgggataatggtctcagaatt 1011
Db 121 ccgagaagaacatggaattgtggtttctcttgggataatggtctcagaatt 180
Qy 1012 tggcggtacactggaacccagacatggaattgtggtttctcttgggataatggtctcagaatt 1071
Db 181 tggcggtacactggaacccagacatggaattgtggtttctcttgggataatggtctcagaatt 240
Qy 1072 ctaccccaaaacatgctgttgcacccagatgacccgtgctttatccacatgctgt 1131
Db 241 ctaccccaaaacatgctgttgcacccagatgacccgtgctttatccacatgctgt 300
Qy 1132 tcccatggtgttatgaaagcatatgcaattggttcccatggtgtgagtcggttctgttt 1191
Db 301 tcccatggtgttatgaaagcatatgcaattggttcccatggtgtgagtcggttctgttt 360
Qy 1192 ggtgatcagatggaacatggaacccagacatggaattgtggtttctcttgggataatggtctcagaatt 1251
Db 361 ggcgatcagatggaacatggaacccagacatggaattgtggtttctcttgggataatggtctcagaatt 420

Qy	2010	aatgatacttaaccaataatcgttcagtcctctcatctctgctctgttctaatagtgcaacc	2069
Db	980	aat-gaacccttaaccaataatcgttcagtcctctcatctctgctctgttctaatagtgcaacc	1039
Qy	2070	tttgtgttttaagaagggaagcttttgtacctttagagtgtag-----	2112
Db	1040	tttgtgttttaagaagggaagcttttgtacctttagagtgtagtaaatgaatgaatgac	1099
Qy	2113	-----	2112
Db	1100	ttggagtgcactgagaacagcatatgattcttcttggggaaaaagaatgatgtatg	1159
Qy	2113	-----gfgaaatgaatgaatgcttggagtcgactggaacagcagcata	2154
Db	1160	aaattgggtgggtgtgtagtgaatgaatgaatgcttggagtcgactggaacagcagcata	1219
Qy	2155	tgattcttctcttgggaaaaagaatgatctatgaattgattgggtgtgtatttgag	2214
Db	1220	tgattcttctcttgggaaaaagaatgatctatgaattgattgggtgtgtatttgag	1279
Qy	2215	aagataataatgcttatgtcaaatggagtcgaattgtgataaaaccccaaatcacgcta	2274
Db	1280	aagataataatgcttatgtcaaatggagtcgaattgtgataaaaccccaaatcacgcta	1339
Qy	2275	tgaagtcgtgggcaagtctactttttctgatcttttccatacaact	2320
Db	1340	tgaagtcgtgggcaagtctactttttctgatcttttccatacaact	1385

RESULT, 10	
AAH51405	
ID	AAH51405 standard; DNA; 1001 BP.
XX	
AC	AAH51405;
XX	
DT	29-AUG-2001 (first entry)
DT	
DE	Human UGT1A7 related DNA containing a biallelic polymorphism SEQ ID 296.
XX	
KW	Human: biallelic marker; single nucleotide polymorphism: SNP; MGS2II;
KW	microsomal glutathione S-transferase II; malate decarboxylase enzyme;
KW	DMET; MEI; cytochrome P450; glutathione reductase; GSHR; GSHS; GGT5;
KW	flavin-containing monooxygenase; FMO; gamma-glutamyltransferase 5;
KW	dispeptidase; DP; glucose 6-phosphate dehydrogenase; G6PDH; haplotype;
KW	phosphogluconate dehydrogenase; PGDH; drug metabolism; phenotype;
KW	uridine diphosphate glucuronosyl transferase; UGT2; asthma; hepatotoxicity;
KW	zileuton; ds.
XX	
OS	homo sapiens.
XX	
PN	WO200058508-A2.
XX	
PD	05-OCT-2000.
XX	
PF	24-MAR-2000; 2000WO-IB00403.
XX	
PR	25-MAR-1999; 99US-0126269.
PR	30-APR-1999; 99US-0131961.
XX	
PA	(GEST) GENSET.
XX	
PI	Blumenfeld M, Bougueleret L, Chumakov I, Cohen-Akenine A;
XX	
DR	WPI; 2000-638353/61.
XX	
PT	Polynucleotides comprising sequences from malate decarboxylase
PT	enzyme-related biallelic markers used for genotyping -
XX	
PS	Claim 13; Page 488-489; 673pp; English.
XX	
CC	Sequences AAH51110-AAH51593 represent human DNA fragments which contain
CC	biallelic markers. The sequences are related to various human genes
CC	including microsomal glutathione S-transferase II (MGS2II), malate

decarboxylase enzyme (DME1/ME1), cytochrome p450, glutathione reductase/synthase (GSHR/GSSH), flavin-containing monooxygenases (FMO), gamma-glutamyltransferase 5 (GGT5), dipeptidase (DP), glucose 6-phosphate dehydrogenase (G6PDH), phosphogluconate dehydrogenase (PGDH), and uridine diphosphate glucuronosyl transferases (UGT2). Each of these sequences contains a biallelic marker/polymorphism, which is represented in the sequence as a degenerate/undefined base. The genes to which the biallelic marker containing sequences are related are involved in drug metabolism. Sequences AAH51594 - AAH51598 represent the genomic sequence of the MGSVII gene and four alternative MGSVII cDNA sequences. AAH62905 - AAH62906 are MGSVII gene products. PCR primers AAH51599 and AAH51600 are used in an example for the amplification of human genomic DNA fragments. The invention includes a method of genotyping comprising determining the identity of a nucleotide at a DME- or MGSVII-related biallelic marker in a biological sample. The method is used to determine the frequency in a population of an allele of a DME- or MGSVII related biallelic marker and to select an individual for inclusion in a clinical trial of a drug treatment. The method is also used to detect association between allele and phenotype, and to detect association between haplotype and phenotype. The polynucleotides are used, in hybridization assays, sequencing assays, or allele specific amplification assays. The method can be used to determine whether an individual suffers or is at risk of developing asthma or is at risk of developing hepatotoxicity on treatment with zileuton.

Seq	Sequence	1001 BP;	315 A;	169 C;	205 G;	311 T;	1 other;
	Query Match	33.2%	Score	769.6;	DB	21;	Length 1001;
	Best Local Similarity	99.9%;	Pred.	No. 3e-208;			
	Matches 769;	Conservative	1;	Mismatches	0;	Indels	0;
QY	1551	ccggaaatccttggggaaaaaaggcgagtttaagaaagcccaaaatccaagaccattg	1610				
Db	1	ccggaaatccttggggaaaaaaggcgagtttaagaaagcccaaaatccaagaccattg	60				
QY	1611	agaagtggttggggaaaaaagtgtaaaatttgaaccattccctcagtcatttccaaacttga	1670				
Db	61	agaagtggttggggaaaaaagtgtaaaatttgaaccattccctcagtcatttccaaacttga	120				
QY	1671	aaacagaatcagtgtaaaattcattttattcttataaggaatacatttgcataataaa	1730				
Db	121	aaacagaatcagtgtaaaattcattttattcttataaggaatacatttgcataataaa	180				
QY	1731	tacgccccagagtgctttaaaaaattctcttaataaaaaataatagactcgtacgt	1790				
Db	181	tacgccccagagtgctttaaaaaattctcttaataaaaaataatagactcgtacgt	240				
QY	1791	aaagatatattgaatatgtatcgtgcacctccggtgtcttgatcagatgacactgtcc	1850				
Db	241	aaagatatattgaatatgtatcgtgcacctccggtgtcttgatcagatgacactgtcc	300				
QY	1851	attttcagagaagctgcagacaggctggaattctatagatacttttactctgaaca	1910				
Db	301	attttcagagaagctgcagacaggctggaattctatagatacttttactctgaaca	360				
QY	1911	tggcctgtttgggagtcgggattcaaaagtggtccacacgctgcctcactgcgaatgg	1970				
Db	361	tggcctgtttgggagtcgggattcaaaagtggtccacacgctgcctcactgcgaatgg	420				
QY	1971	cagttttaattctatcttttggctctgcgaatggttgcaatgatccttaaccaataat	2030				
Db	421	cagttttaattctatcttttggctctgcgaatggttgcaatgatccttaaccaataat	480				
QY	2031	ggtcagtcctcatctcgtctcttcattaggtgccaacctgtgtgtttaagaaggaa	2090				
Db	481	ggtcagtcctcatctcgtctcttcattaggtgccaacctgtgtgtttaagaaggaa	540				
QY	2091	gctttgtacotttagatgtagtgaaatgaatgaatggtctggagtgcactgaaacag	2150				
Db	541	gctttgtacotttagatgtagtgaaatgaatgaatggtctggagtgcactgaaacag	600				
QY	2151	cataataattcttgcttggagaaagaaagaaatgactgatgaaattgctgggtgtgatt	2210				

QY 1891 actttcttactctgaacatgacctgttttggagtcgaggattcaaaagtggtccacc 1950
 Db 15167 actttcttactctgaacatgacctgttttggagtcgaggattcaaaagtggtccacc 15226
 QY 1951 gctgcccactactgcaaatgacgttttaactcttactcttttggcttctgcagatggttga 2010
 Db 15227 gctgcccactactgcaaatgacgttttaactcttactcttttggcttctgcagatggttga 15286
 QY 2011 attgctcttaacaaataatggtcagtcctcactctctgtctctcttcataggtgccact 2070
 Db 15287 attgctcttaacaaataatggtcagtcctcactctctgtctctcttcataggtgccact 15346
 QY 2071 tgtgtgttttaaaagaaggaagctttgtaccttttagagtgagtgaaatgaatggc 2130
 Db 15347 tgtgtgttttaaaagaaggaagctttgtaccttttagagtgagtgaaatgaatggc 15406
 QY 2131 tggagtgactgagacagcatatgattcttctgtcttggggaagaaatgatgctatg 2190
 Db 15407 tggagtgactgagacagcatatgattcttctgtcttggggaagaaatgatgctatg 15466
 QY 2191 aaattgggtgggtgtgtatttgaagaataatcattgtcttattgcaaatggagctgaatt 2250
 Db 15467 aaattgggtgggtgtgtatttgaagaataatcattgtcttattgcaaatggagctgaatt 15526
 QY 2251 tgaataaaacccaaatacagcatatgaagtgctgggcaagtttacttttttctgatgtt 2310
 Db 15527 tgaataaaacccaaatacagcatatgaagtgctgggcaagtttacttttttctgatgtt 15586
 QY 2311 tctacaact 2320
 Db 15587 tctacaact 15596

RESULT 7

AAH51403

ID AAH51403 standard; DNA; 1001 BP.

AC AAH51403;

DT 29-AUG-2001 (first entry)

DE Human UGT1A7 related DNA containing a bi-allelic polymorphism SEQ ID 294.

XX Human; bi-allelic marker; single nucleotide polymorphism; SNP; MGSII;
 KW microsomal glutathione S-transferase II; malate decarboxylase enzyme;
 KW DME1; ME1; cytochrome P450; glutathione reductase; GSHR; GSHS; GGT5;
 KW flavin-containing monooxygenase; FMO; gamma-glutamyltransferase 5;
 KW dipeptidase; DP; glucose 6-phosphate dehydrogenase; G6PDH; haplotype;
 KW phosphogluconate dehydrogenase; PGDH; drug metabolism; phenotype;
 KW uridine diphosphate glucuronosyl transferase; UGT2; asthma; hepatotoxicity;
 KW zileuton; ds.

XX Homo sapiens.

XX WO200058508-A2.

XX 05-OCT-2000.

XX 24-MAR-2000; 2000WO-IB00403.

XX 25-MAR-1999; 99US-0126269.

XX 30-APR-1999; 99US-0131961.

XX (GEST) GENSET.

XX Blumenfeld M, Bougueleret L, Chumakov I, Cohen-Akenine A;

XX WPI; 2000-638353/61.

XX Polynucleotides comprising sequences from malate decarboxylase
 PT enzyme-related bi-allelic markers used for genotyping -

XX Claim 13; Page 487; 673pp; English.

XX

CC Sequences AAH51110-AAH51593 represent human DNA fragments which contain
 CC bi-allelic markers. The sequences are related to various human genes
 CC including microsomal glutathione S-transferase II (MGSII), malate
 CC decarboxylase enzyme (DME1/ME1), cytochrome P450, glutathione
 CC reductase/synthase (GSHR/GSHS), flavin-containing monooxygenases (FMO),
 CC gamma-glutamyltransferase 5 (GGT5), dipeptidase (DP), glucose 6-phosphate
 CC dehydrogenase (G6PDH), phosphogluconate dehydrogenase (PGDH), and uridine
 CC diphosphate glucuronosyl transferases (UGT2). Each of these sequences
 CC contains a bi-allelic marker/polymorphism, which is represented in the
 CC sequence as a degenerate/undefined base. The genes to which the bi-allelic
 CC marker containing sequences are related are involved in drug metabolism.
 CC Sequences AAH51594 - AAH51598 represent the genomic sequence of the
 CC MGSII gene and four alternative MGSII cDNA sequences. AAH62905-AAH62906
 CC are MGSII gene products. PCR primers AAH51599 and AAH51600 are used in
 CC an example for the amplification of human genomic DNA fragments. The
 CC invention includes a method of genotyping comprising determining the
 CC identity of a nucleotide at a DME- or MGSII-related bi-allelic marker in
 CC a biological sample. The method is used to determine the frequency in
 CC population of an allele of a DME- or MGSII-related bi-allelic marker and
 CC to select an individual for inclusion in a clinical trial of a drug
 CC treatment. The method is also used to detect association between allele
 CC and phenotype, and to detect association between haplotype and phenotype.
 CC The polynucleotides are used, in hybridization assays, sequencing assays,
 CC or allele specific amplification assays. The method can be used to
 CC determine whether an individual suffers or is at risk of developing
 CC asthma or is at risk of developing hepatotoxicity on treatment with
 CC zileuton.

XX

SQ Sequence 1001 BP; 269 A; 205 C; 233 G; 293 T; 1 other;

Query Match

43.0%; Score 998.6; DB 21; Length 1001;

Best Local Similarity 99.9%; Pred. No. 2.1e-273;

Matches 998; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1322 agaacatcatgcctctccagccttcacaaagacgcccgggtggagccgctggacctgg 1381

Db 1 agaacatcatgcctctccagccttcacaaagacgcccgggtggagccgctggacctgg 60

QY 1382 ccgtgtcttgggtggagcttctgtagggcacaagggcgccacacccctgcgcccgcaag 1441

Db 61 ccgtgtcttgggtggagcttctgtagggcacaagggcgccacacccctgcgcccgcaag 120

QY 1442 cccagcactcactcctgaccagaccattccttgagcagtgatggttctctctggccg 1501

Db 121 cccagcactcactcctgaccagaccattccttgagcagtgatggttctctctggccg 180

QY 1502 tctgtctgacagtgcccttcacacctttaaattgtgtgcttggctaccggaatgct 1561

Db 181 tctgtctgacagtgcccttcacacctttaaattgtgtgcttggctaccggaatgct 240

QY 1562 tggggaaaaaaggcgagtttaagaaagcccaataatccaaagccattgagaagtggtg 1621

Db 241 tggggaaaaaaggcgagtttaagaaagcccaataatccaaagccattgagaagtggtg 300

QY 1622 ggaataaagggtaaaattttgaaccattcccttagctattcccaacttgaacacagaatca 1681

Db 301 ggaataaagggtaaaattttgaaccattcccttagctattcccaacttgaacacagaatca 360

QY 1682 gtgttaataatcatttatttatttaagaaataactttgcataataataatcagcccca 1741

Db 361 gtgttaataatcatttatttatttaagaaataactttgcataataataatcagcccca 420

QY 1742 gtgtttaaaaaattctctttaaataataaataactcgtcagtaagaatatttg 1801

Db 421 gtgtttaaaaaattctctttaaataataaataactcgtcagtaagaatatttg 480

QY 1802 aatgtatctggtccctccgttcttcttgatcagtgatgagtgatgctatttttcagag 1861

Db 481 aatgtatctggtccctccgttcttcttgatcagtgatgagtgatgctatttttcagag 540

QY 1862 gacgtgcagacaggtggcattcttagattactttcttacttctgaaacatggcctgtttg 1921

CC to restore normal activity of (II) or to treat disease states involving
 CC (III). (II) is useful for generating antibodies against it, detecting or
 CC quantitating a polypeptide in tissue, as molecular weight markers and as
 CC a food supplement. (II) and its binding partners are useful in medical
 CC imaging of sites expressing (II). (I) and (II) are useful for treating
 CC disorders involving aberrant protein expression or biological activity.
 CC The polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. AAS64197-AAS94564 represent novel human
 CC diagnostic coding sequences of the invention.
 CC Note: The sequence data for this patent did not appear in the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 XX
 SQ Sequence 1655 BP; 402 A; 388 C; 403 G; 462 T; 0 other;

Query Match 66.9%; Score 1551.8; DB 23; Length 1655;
 Best Local Similarity 97.7%; Pred. No. 0;
 Matches 1574; Conservative 0; Mismatches 37; Indels 0; Gaps 0;

QY 1 ctcaagctgcagttctctgctgacaggtgacaggtgacagcccccttctctatgtgtg 60
 DB 45 ctcaagctgcagttctctgctgacaggtgacaggtgacagcccccttctctatgtgtg 104
 QY 61 tctc-gctgtgacctgtggtttgcccagggaggaagctactggtagtgcacatggt 120
 DB 105 tctc-gctgtgacctgtggtttgcccagggaggaagctactggtagtgcacatggt 164
 QY 121 gggagccactggttcacatgagtgctggtgagaaactcattctcaggggacatag 180
 DB 165 gggagccactggttcacatgagtgctggtgagaaactcattctcaggggacatag 224
 QY 181 gtggtgtgtagtcgacagaggtgagttggcaactgggaagatcactgaattgcacagt 240
 DB 225 gtggtgtgtagtcgacagaggtgagttggcaactgggaagatcactgaattgcacagt 284
 QY 241 aagacttatcacaactcattacacctggaggtgctggccgggagttcgaagctttgcc 300
 DB 285 aagacttatcacaactcattacacctggaggtgctggccgggagttcgaagctttgcc 344
 QY 301 catgtcaaatggaagacagtagtcagaggtatattctctatattatggtttcacaat 360
 DB 345 catgtcaaatggaagacagtagtcagaggtatattctctatattatggtttcacaat 404
 QY 361 gacatatttgacttatcttttcaaatgacaggtgttggtaagacaaaataatgta 420
 DB 405 ggttttttaacttatttttcgattgcagaggtgttggtaagacaaaataatgta 464
 QY 421 gaatactaaaggagaggtcttttgatgcaggtgttctcgatcccttttgataactgtggc 480
 DB 465 gaatactaaaggagaggtcttttgatgcaggtgttctcgatcccttttgatgcaggtgtggc 524
 QY 481 ttaattgtgtccaaatatttctccctccctcctggtgttctgcaggggaatacttgc 540
 DB 525 ttaattgtgtccaaatatttctccctccctcctggtgttctgcaggggaatacttgc 584
 QY 541 cactatctgaagaggtgacagtgccctgtctctcttctctatgtccccagaattctc 600
 DB 585 cactatctgaagaggtgacagtgccctgtctctcttctctatgtccccagaattctc 644
 QY 601 ttagggttctcagatgcacatgactttcaaggagagtagcaggaacacatcatcacttg 660
 DB 645 ttagggttctcagatgcacatgactttcaaggagagtagcaggaacacatcatcacttg 704
 QY 661 gaggaaacttattatgcacaggttttttcaaaaatgccttagaataagcctctgaaatt 720
 DB 705 gaggaaacttattatgcacaggttttttcaaaaatgccttagaataagcctctgaaatt 764
 QY 721 ctccaaacacctgttacggagtagtctctcagccacacatcaatttggtgttgga 780
 DB 1655 ctccaaacacctgttacggagtagtctctcagccacacatcaatttggtgttgga 824

DB 765 ctccaaacacctgttacggagtagtctctcagccacacatcaatttggtgttgga 824
 QY 781 acgagactttgtttgactatcccaaacccgtgatgcccaacatgatcttcattggtgt 840
 DB 825 acagactttgtttgactatcccaaacccgtgatgcccaacatgatcttcattggtgtgt 884
 QY 841 atcaactgcatcagggaaagccgttccctatggaattggaagcctacataatgtctct 900
 DB 885 atcaactgcatcagggaaagccattgctatggaattggaagcctacataatgtctct 944
 QY 901 ggagaaacatgaaattgtgtttcttcttgggataatgttctcagaaattccagaaag 960
 DB 945 ggagaaacatgaaattgtgtttcttcttgggataatgttctcagaaattccagaaag 1004
 QY 961 aaagctatgcaattgctgagttgttgggcaaaatccctcagacagtcctgtggcggtac 1020
 DB 1005 aaagctatgcaattgctgagttgttgggcaaaatccctcagacagtcctgtggcggtac 1064
 QY 1021 actggaaccccgacatcgaatcttgcgaacacacagactactgttaagtgcctaccacaa 1080
 DB 1065 actggaaccccgacatcgaatcttgcgaacacacagactactgttaagtgcctaccacaa 1124
 QY 1081 aacgactgtgttgcaccccgatgacccgtgcttcttatacccatgctgttcccatggt 1140
 DB 1125 aacgactgtgttgcaccccgatgacccgtgcttcttatacccatgctgttcccatggt 1184
 QY 1141 gtttatgaaagcatalgcaatggcgttcccatggtgatgacccctgttctgtgatcag 1200
 DB 1185 gtttatgaaagcatalgcaatggcgttcccatggtgatgacccctgttctgtgatcag 1244
 QY 1201 atggacaatgcaaaagcagatgagactaaaggagagtgagtgacccctgaatgtcttgaa 1260
 DB 1245 atggacaatgcaaaagcagatgagactaaaggagagtgagtgacccctgaatgtcttgaa 1304
 QY 1261 atgacttctgaaattgaaatgctctaaaggcagtcacatgacaaagattacaag 1320
 DB 1305 atgacttctgaaattgaaatgctctaaaggcagtcacatgacaaagattacaag 1364
 QY 1321 gagaacatcagcctcagcctcagcctcagcctcagcctcagcctcagcctcagcctg 1380
 DB 1365 gagaacatcagcctcagcctcagcctcagcctcagcctcagcctcagcctcagcctg 1424
 QY 1381 gccgtgttctgggtggagttgtgatgagcacaaggcgcgcacacacctgcgcccgcga 1440
 DB 1425 gccgtgttctgggtggagttgtgatgagcacaaggcgcgcacacacctgcgcccgcga 1484
 QY 1441 gccac 1500
 DB 1485 gccac 1544
 QY 1501 gtcgtgctgacagtgcccttccatcaccttttaaatgttgccttatggctaccggaaatgc 1560
 DB 1545 gtcgtgctgacagtgcccttccatcaccttttaaatgttgccttatggctaccggaaatgc 1604
 QY 1561 ttggggaaaaagggcgagtttaagaaagcccaacacacacacacacacacacacacacac 1611
 DB 1605 ttggggaaaaagggcgagtttaagaaagcccaacacacacacacacacacacacacacac 1655

RESULT 6

AAS18543

ID AAS18543 standard; DNA; 18887 BP.

XX AAS18543;

XX AAS18543;

DT 12-MAR-2002 (first entry)

XX DNA encoding UDP glycosyltransferase 1 (UGT1A1).

DE UDP glycosyltransferase 1; UGT1A1; human; haplotyping;
 KW drug discovery; Gilbert's syndrome; Crigler-Najjar syndrome; ds.

XX Homo sapiens.

OS

QY 858 aaagccgttgctatggaatttgaagcctacattaatgtcttctggagaacatggaattgt 917
Db 881 ggaagccactatcaggaatttgaagcctacattaatgtcttctggagaacatggaattgt 940
QY 918 ggtttctcttggatcgaatttgaagcctacattaatgtcttctggagaacatggaattgt 977
Db 941 ggtttctcttggatcgaatttgaagcctacattaatgtcttctggagaacatggaattgt 1000
QY 978 tgaatgcttgggaacacccctcagacagtgcttggcgtgacatggaacccgacacac 1037
Db 1001 tgaatgcttgggaacacccctcagacagtgcttggcgtgacatggaacccgacacac 1060
QY 1038 gaattctgggaac 1097
Db 1061 gaattctgggaac 1120
QY 1098 ccgatgacccgtgcttcttcaacacacacacacacacacacacacacacacacacac 1157
Db 1121 ccgatgacccgtgcttcttcaacacacacacacacacacacacacacacacacacac 1180
QY 1158 caatggcgttcccatggtgatgacacacacacacacacacacacacacacacacacac 1217
Db 1181 caatggcgttcccatggtgatgacacacacacacacacacacacacacacacacacac 1240
QY 1218 catgagac 1277
Db 1241 catgagac 1300
QY 1278 agaaaatgcttcaaaagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1337
Db 1301 agaaaatgcttcaaaagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1360
QY 1338 ctccagcctcacaagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1397
Db 1361 ctccagcctcacaagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1420
QY 1398 gttgtgatgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1457
Db 1421 gttgtgatgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1480
QY 1458 gtaccagtcac 1517
Db 1481 gtaccagtcac 1540
QY 1518 ctcatcac 1577
Db 1541 ctcatcac 1600
QY 1578 agtgaagaagccac 1637
Db 1601 agtgaagaagccac 1660
QY 1638 ttgaac 1697
Db 1661 ttgaac 1720
QY 1698 attcttattgaagaac 1757
Db 1721 attcttattgaagaac 1779
QY 1758 tcttaataataataataataataataataataataataataataataataataataata 1817
Db 1780 tcttaataataataataataataataataataataataataataataataataataata 1839
QY 1818 cctccggtgtcttgcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1877
Db 1840 cctccggtgtcttgcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1899
QY 1878 ggcattctagattcttcttcttcttcttcttcttcttcttcttcttcttcttcttctt 1937
Db 1900 ggcattctagattcttcttcttcttcttcttcttcttcttcttcttcttcttcttctt 1959

QY 1938 agtgatgctccacccgctgcccctactgcaaatggcagtttatacttcttcttcttcttct 1997
Db 1960 agtgatgctccacccgctgcccctactgcaaatggcagtttatacttcttcttcttcttct 2019
QY 1998 gcagatgcttgcgaattgaccccttaacccaataatggcagtcctcctcctcctcctcct 2057
Db 2020 gcagatgcttgcgaattgaccccttaacccaataatggcagtcctcctcctcctcctcct 2079
QY 2058 atagtgccacc-ttgtgttttaaaaggaagccttcttcttcttcttcttcttcttcttct 2116
Db 2080 atagtgccacc-ttgtgttttaaaaggaagccttcttcttcttcttcttcttcttcttct 2139
QY 2117 aatgaatgaatggcttggagtcactgagacacacacacacacacacacacacacacac 2176
Db 2140 aatgaatgaatggcttggagtcactgagacacacacacacacacacacacacacacac 2199
QY 2177 aagaatgctatgaatggcttggagtcactgagacacacacacacacacacacacacacac 2236
Db 2200 aagaatgctatgaatggcttggagtcactgagacacacacacacacacacacacacacac 2259
QY 2237 aatgagcgtgaatttataaaac 2296
Db 2260 aatgagcgtgaatttataaaac 2319
QY 2297 tttttctgagtttctcacaact 2320
Db 2320 tttttctgagtttctcacaact 2343

RESULT 5
AAS91252
ID AAS91252 standard; cDNA; 1655 BP.
XX AAS91252;
XX
DT 13-FEB-2002 (first entry)
XX
XX DNA encoding novel human diagnostic protein #27056.
XX Human; chromosome mapping; gene mapping; gene therapy; forensic;
XX food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX Homo sapiens.
XX WO200175067-A2.
XX 11-OCT-2001.
XX 30-MAR-2001; 2001WO-US08631.
XX 31-MAR-2000; 2000US-0540217.
XX 23-AUG-2000; 2000US-0649167.
XX (HYSE-) HYSEQ INC.
XX Drmanac RT, Liu C, Tang YT;
XX WPI; 2001-639362/73.
XX P-PSDB; AB627065.
XX
XX New isolated polynucleotide and encoded polypeptides, useful in
XX diagnostics, forensics, gene mapping, identification of mutations
XX responsible for genetic disorders or other traits and to assess
XX biodiversity -
XX
XX Claim 1; SEQ ID No 27056; 103pp; English.
XX
XX The invention relates to isolated polynucleotide (I) and
XX polypeptide (II) sequences. (I) is useful as hybridisation probes,
XX polymerase chain reaction (PCR) primers, oligomers, and for chromosome
XX mapping, and in recombinant production of (II). The
XX polynucleotides are also used in diagnostics as expressed sequence tags
XX for identifying expressed genes. (I) is useful in gene therapy techniques

Db 1870 gtgcagacagcgctggtcattctagattacttttttactctgaaacatgacctgtttggga 1929
 Qy 1925 gtgcgggattcaagggtgtccaccgctgcccactactgcaaatgagtttttaacttta 1984
 Db 1930 gtgcgggattcaagggtgtccaccgctgcccactactgcaaatgagtttttaacttta 1989
 Qy 1985 tcttttgctctgcagatggttgcattgacttcaaccataatggtcagtcctcatc 2044
 Db 1990 tcttttgctctgcagatggttgcattgacttcaaccataatggtcagtcctcatc 2049
 Qy 2045 tctgtcctgcttcatagggtccacc-ttgtgtgttttaagaagggaagcctttgtactctt 2103
 Db 2050 tctgtcctgcttcatagggtccacccttgtgtgttttaagaagggaagcctttgtactctt 2109
 Qy 2104 agagtgtaggtgaaatgaatgaatggttgcattgactgcaacacacatgattctt 2163
 Db 2110 agagtgtaggtgaaatgaatgaatggttgcattgactgcaacacacatgattctt 2169
 Qy 2164 gctttggggaaaaagatgctatgataattggtgtgtgtatttgagaagataatc 2223
 Db 2170 gctttggggaaaaagatgctatgataattggtgtgtgtatttgagaagataatc 2229
 Qy 2224 attgcttatgtcgaatgagctgaatttgataaaaccccaaaacacagctatgaagtct 2283
 Db 2230 attgcttatgtcgaatgagctgaatttgataaaaccccaaaacacagctatgaagtct 2289
 Qy 2284 gggcnaagttacttttttctgattgttctcacaact 2320
 Db 2290 gggcnaagttacttttttctgattgttctcacaact 2326

RESULT 4

AAQ27370
 ID AAQ27370 standard; cDNA; 2368 BP.
 AC AAQ27370;
 XX
 XX
 XX 27-JAN-1993 (first entry)
 DE HUG-Br2.
 XX
 KW Billirubin; UDP-glucuronosyltransferase; HUGBr1; HUGBr2;
 KW monoglucuronide; diglucuronide; ss.
 XX Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT CDS 30..801
 FT /tag= a
 FT polyA_signal 2347..2352
 FT /tag= b
 FT /number= 1
 FT polyA_signal 2355..2360
 FT /tag= c
 FT /number= 2
 XX
 PN W09212987-A.
 XX
 PD 06-AUG-1992.
 XX
 PF 10-JAN-1992; 92WO-US00282.
 XX
 PR 10-JAN-1991; 91US-0639453.
 XX
 PA (USSH) US DEPT HEALTH & HUMAN SERVICE.
 XX
 PI Owens IS, Ritter JK;
 XX
 DR WPI: 1992-284593/34.
 DR P-PSDB: AAR26154.
 XX
 XX Isolated gene locus UGRI, DNA segments and diagnostic probes -
 PT for diagnosing Gilbert's disease and Crigler-Najjar syndrome

types I and II

XX
 PS Disclosure; Fig 9A-I; 99pp; English.
 XX
 CC Two human liver bilirubin UDP-glucuronosyltransferase cDNAs have
 CC been isolated. They are referred to as HUGBr1 (AAQ27369) and HUGBr2
 CC (AAQ27370) (Ritter, et al., J. Biol. Chem. 266:1043-1047 (1991)) and,
 CC upon expression individually in COS-1 cells, encode isoforms that
 CC catalyze the formation of the two bilirubin monoglucuronides and
 CC the diglucuronide.
 CC The cDNAs contain identical 3' ends (1469 bp in length) to each
 CC other and to that of the human phenol transferase cDNA, HUGP1
 CC (Harding et al., Proc. Natl. Acad. Sci. USA 85:8281 (1988)).
 CC In contrast, they have unique 5' ends.
 XX
 SQ Sequence 2368 BP; 609 A; 532 C; 566 G; 661 T; 0 other;

Query Match 68.0%; Score 1576.8; DB 13; Length 2368;
 Best Local Similarity 82.3%; Pred. No. 0;
 Matches 1847; Conservative 0; Mismatches 392; Indels 5; Gaps 3;
 QY 81 ctttgcgagcgagggaagctactgttagtgcctatggagcagcctgttccacct 140
 Db 101 ctgggctgagagtggaaaggtgtgtgtggtgcccactgtgagccctgttccatg 160
 QY 141 gaggtcggtgtgtgagaacactcttcaggggcatgaggtgtgttagtcatgccaga 200
 Db 161 gcgggagggccttgcggagctccatgccagagccaccagcgggtgctcaccaccaag 220
 QY 201 ggtgagttggcaactgggaagatcaactgaatgcacagtgaaagactttcaactcata 260
 Db 221 aggtgaatatgcacatcaagaagaataattttccctgcacgcctatgctgtccatg 280
 QY 261 taccctggaggtatgcacggggagttcaagggttttgcctatgctcaat---ggaaagc 317
 Db 281 gaccacagaagaaatttgatcggttacgctgggtcacactcaagggttttgaacacaga 340
 QY 318 acaagtacgaagtatatattctctatttaattggtgtcacaatgaacatttttgactttt 377
 Db 341 acatctctgaagagatattctagaagtaaggcaattatgaacattgtatctttggccct 400
 QY 378 ttttcaaatgacagaggtttgtttaaagacaaaataattagataacttaaaaggagag 437
 Db 401 tcataggtgtgtgtgagctactgataatgagccctgacagcagcactgaatgctac 460
 QY 438 tcttttgatgcaggtttctcgatctcttttgaataactgtggttaattgttgcaata 497
 Db 461 ttcctttgagtggttttaacagaccccttaacctctgcggggcgggtgctggcctaagta 520
 QY 498 tttctcctccctccgtgtgtgttcctgcaggggaataactttgccaactatcttgaagaagg 557
 Db 521 cctgctgattcctgctgtgtttttttggaggtacatttccatgtgacttagactttaagg 580
 QY 558 tgcacagtgctcctctcttctctatgtccccaagaattctctttaggtttctcagatgc 617
 Db 581 cacacagtgctcaaatctctctctataattcttaagttaactaacacacacattcagacca 640
 QY 618 catgactttcaaggagagtagtaagcaacacatcatgcaacttggaggaacatttattatg 677
 Db 641 catgacattctgcgaagggtcaagaacatgctctacacctgtgacctgtctctacatttg 700
 QY 678 ccaccgtttttcaaaaatgccctagaaatagctctctgaaattctcacaacacctgttac 737
 Db 701 ccatacttttctgccccttatgcaagctctgtcctctgagcgttttccagagagaggtagc 760
 QY 738 ggagtagtatctctacagccacacataattgtgtgtgcgaacgcagctttgttttga 797
 Db 761 tgtccaggaaactattgactctgcatctgtctggtgttttagaagtgactttgtgaagga 820
 QY 798 ctatcccaacccgtgagcccaacatgcttctggtgtatcactgacactgccatcagg 857
 Db 821 ttaccttaggccctcatcgtcccaataatggtcttcttcttggggggcatcaactgttgcacac 880

QY 781 aaggactttgttttgactatccccaaacccgtgatgccaaacatgatcttcattgtgtgt 840
 Db 781 aaggactttgttttgactatccccaaacccgtgatgccaaacatgatcttcattgtgtgt 840
 QY 841 atcaactgccatcagggaagccgttgctctatgaatttgaagccacatttaattgtct 900
 Db 841 atcaactgccatcagggaagccgttgctctatgaatttgaagccacatttaattgtct 900
 QY 901 gagaaacatgaattgtgttttctcttgggatacaatggtctcagaataatccagagaag 960
 Db 901 gagaaacatgaattgtgttttctcttgggatacaatggtctcagaataatccagagaag 960
 QY 961 aagctatggcaattgctgatgctttgggcaaaatccctcagacagctcgtggcggtac 1020
 Db 961 aagctatggcaattgctgatgctttgggcaaaatccctcagacagctcgtggcggtac 1020
 QY 1021 actggaacccgacacatcgaattctgcgaacacacacatacttgttaagtggctaccocaa 1080
 Db 1021 actggaacccgacacatcgaattctgcgaacacacacatacttgttaagtggctaccocaa 1080
 QY 1081 aacgatctgcttgggtcacccgatgacccgtgacctttatcacccatgctgttcccatagt 1140
 Db 1081 aacgatctgcttgggtcacccgatgacccgtgacctttatcacccatgctgttcccatagt 1140
 QY 1141 gtttatgaagcatacgaattggcggttcccatggtgatgacccctgttgggtgatcag 1200
 Db 1141 gtttatgaagcatacgaattggcggttcccatggtgatgacccctgttgggtgatcag 1200
 QY 1201 atggnaaatcgaagcgcatggagactaaaggagctggagtgacccctgaatgtcttgaa 1260
 Db 1201 atggnaaatcgaagcgcatggagactaaaggagctggagtgacccctgaatgtcttgaa 1260
 QY 1261 atgactctgaagatttgaataatgctctaaaggagctgaatgaatgaatgaatgaatgaat 1320
 Db 1261 atgactctgaagatttgaataatgctctaaaggagctgaatgaatgaatgaatgaatgaat 1320
 QY 1321 gagaacatcgtgcctctccagcttccaaagacccgctggagccgctggagccgctggagccg 1380
 Db 1321 gagaacatcgtgcctctccagcttccaaagacccgctggagccgctggagccgctggagccg 1380
 QY 1381 gccgtgttgagtggaattgtgtaggacacagggcgccacacactgcgccccga 1440
 Db 1381 gccgtgttgagtggaattgtgtaggacacagggcgccacacactgcgccccga 1440
 QY 1441 gccac 1500
 Db 1441 gccac 1500
 QY 1501 gctgtgctgacagtgccctcaccctttaaattgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1560
 Db 1501 gctgtgctgacagtgccctcaccctttaaattgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1560
 QY 1561 ttggcgaaaaagggcgagtttaagaaagcccaatcccaagaccccatgtgagaagtggt 1620
 Db 1561 ttggcgaaaaagggcgagtttaagaaagcccaatcccaagaccccatgtgagaagtggt 1620
 QY 1621 gggaaataaggttaaaatttgaacacatccctagtcatttccaaactgaaacagaatc 1680
 Db 1621 gggaaataaggttaaaatttgaacacatccctagtcatttccaaactgaaacagaatc 1680
 QY 1681 agtgttaatttattttattttattttattttattttattttattttattttattttattttatttt 1740
 Db 1681 agtgttaatttattttattttattttattttattttattttattttattttattttattttatttt 1740
 QY 1741 agtgttttaaaaaattctttaaataaaaaataatagactcgtagtcagtaagaatattt 1800
 Db 1741 agtgttttaaaaaattctttaaataaaaaataatagactcgtagtcagtaagaatattt 1800
 QY 1801 gaatagtatcgtgccctccgtgtcttcttgatcaggatgacatgtgccatttttcaga 1860
 Db 1801 gaatagtatcgtgccctccgtgtcttcttgatcaggatgacatgtgccatttttcaga 1860

QY 1861 ggacgtgcagacagcgtggcattcttagattacttttttacttactgaaacatggccgtttt 1920
 Db 1861 ggacgtgcagacagcgtggcattcttagattacttttttacttactgaaacatggccgtttt 1920
 QY 1921 gggagtcgggattcgaaggtgtgtccaccgctgcccctactgcaaatggcagttttaa 1980
 Db 1921 gggagtcgggattcgaaggtgtgtccaccgctgcccctactgcaaatggcagttttaa 1980
 QY 1981 cttaacttttggcttctgcagatggttgcaatgcatcttaaccaataatggctcagctct 2040
 Db 1981 cttaacttttggcttctgcagatggttgcaatgcatcttaaccaataatggctcagctct 2040
 QY 2041 catctctgctctctcataggtggccacctgtgtgtttaaagaagggaagcctttgtacc 2100
 Db 2041 catctctgctctctcataggtggccacctgtgtgtttaaagaagggaagcctttgtacc 2100
 QY 2101 tttagagtgtagtgtaaatgaatgaatgagcttgagctgcaatgagcagacatgatgatt 2160
 Db 2101 tttagagtgtagtgtaaatgaatgaatgagcttgagctgcaatgagcagacatgatgatt 2160
 QY 2161 ctgtctttggggaaaaagaatgagctgcatgaaatggtgtgtgtgtgtgtgtgtgtgtgtgtgt 2220
 Db 2161 ctgtctttggggaaaaagaatgagctgcatgaaatggtgtgtgtgtgtgtgtgtgtgtgtgtgt 2220
 QY 2221 atcattgcttctgcaaatggagctgaattgtataaaaccccaaaacacagctatgaagt 2280
 Db 2221 atcattgcttctgcaaatggagctgaattgtataaaaccccaaaacacagctatgaagt 2280
 QY 2281 gctgggcaaatcttaacttttttctgatttctcacaact 2320
 Db 2281 gctgggcaagttaacttttttctgatttctcacaact 2320

RESULT 3
 AAQ27369
 ID AAQ27369 standard; cDNA; 2351 BP.
 XX
 AC AAQ27369;
 XX
 DT 27-JAN-1993 (first entry)
 XX
 DE HUG-Br1.
 XX
 KW Bilirubin; UDP-glucuronosyltransferase; HUGBr1; HUGBr2;
 KW monoglucuronide; diglucuronide; ss.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT CDS 16..784
 FT /*tag= a
 FT polyA_signal 2330..2335
 FT /*tag= b
 FT /*number= 1
 FT polyA_signal 2338..2343
 FT /*tag= c
 FT /*number= 2
 XX
 PN W09212987-A.
 XX
 PD 06-AUG-1992.
 XX
 PF 10-JAN-1992; 92MO-US00282.
 XX
 PR 10-JAN-1991; 91US-0639453.
 XX
 PA (USSH) US DEP. HEALTH & HUMAN SERVICE.
 XX
 PI Owens IS, Ritter JK;
 XX
 DR WPI; 1992-284593/34.
 DR P-PSDB; AAR26153.
 XX

Qy 1801 gaattgtatgctgccccctccggtgttcttggatcaggatgacatgcatgtgcccatttttcaga 1860
 |||||
 Db 1801 gaattgtatgctgccccctccggtgttcttggatcaggatgacatgcatgtgcccatttttcaga 1860
 |||||
 Qy 1861 ggaactgcaacagagctgcatctctagattactttttctactctgaaacatgcccgttt 1920
 |||||
 Db 1861 ggaactgcaacagagctgcatctctagattactttttctactctgaaacatgcccgttt 1920
 |||||
 Qy 1921 gggagtcgggattcaaaagtgtcccccacgcgtccctactgcaaatggcagtttaatt 1980
 |||||
 Db 1921 gggagtcgggattcaaaagtgtcccccacgcgtccctactgcaaatggcagtttaatt 1980
 |||||
 Qy 1981 cttatcttttgccttgcagatgttgcattgacatccttaaccaataatggcagtcct 2040
 |||||
 Db 1981 cttatcttttgccttgcagatgttgcattgacatccttaaccaataatggcagtcct 2040
 |||||
 Qy 2041 catctctgctcttctatagtgccacaccttgtgttttaaaagaggaagctttgtacc 2100
 |||||
 Db 2041 catctctgctcttctatagtgccacaccttgtgttttaaaagaggaagctttgtacc 2100
 |||||
 Qy 2101 tttagagtgtagtgaaatgaatgaatgcttgagtgactgagacagacatgatatt 2160
 |||||
 Db 2101 tttagagtgtagtgaaatgaatgaatgcttgagtgactgagacagacatgatatt 2160
 |||||
 Qy 2161 cttgctttggggaaaaagatgatgctatgaaattggtgggtgtatttgagaagata 2220
 |||||
 Db 2161 cttgctttggggaaaaagatgatgctatgaaattggtgggtgtatttgagaagata 2220
 |||||
 Qy 2221 atcattgcttatgtcaaatggagctgaatttgataaaaccccaaaatacagctatgaat 2280
 |||||
 Db 2221 atcattgcttatgtcaaatggagctgaatttgataaaaccccaaaatacagctatgaat 2280
 |||||
 Qy 2281 gctgggcaagtttactttttctctgatgttctctacaact 2320
 |||||
 Db 2281 gctgggcaagtttactttttctctgatgttctctacaact 2320
 |||||

RESULT 2

ABA05199 standard; cDNA; 2320 BP.
 ID ABA05199 standard; cDNA; 2320 BP.
 XX
 AC ABA05199;
 XX
 XX
 DT 04-MAR-2002 (first entry)
 XX
 DE Human uridine 5'diphosphate glucuronyltransferase UGT1A9 cDNA #2.
 XX
 KW Human; uridine 5'diphosphate glucuronyltransferase enzyme; UGT1A9;
 KW flavopiridol; cancer; gastrointestinal disease; parasitic infection;
 KW cytosolic; ss.
 XX
 OS Homo sapiens.
 XX
 FH key Location/Qualifiers
 FT CDS 19..1611
 FT /*tag= a
 FT /*product= "UGT1A9"
 XX
 PN WO200180896-A2.
 XX
 XX
 PD 01-NOV-2001.
 XX
 PF 12-APR-2001; 2001WO-US12526.
 XX
 PR 21-APR-2000; 2000US-0553829.
 XX
 XX (ARCH-) ARCH DEV CORP.
 PA
 PI Ratain MJ, Innocenti F, Iyer L;
 XX
 DR WPI; 2002-075093/10.
 DR P-PSDB; ABB04430.
 XX

Combinations of flavopiridol and an agent that increases conjugative enzyme activity or glucuronyltransferase activity, with reduced side effects, for treating cancer -

Claim 81; Page 141-144; 145pp: English.

The present invention relates to a method of reducing the toxicity of flavopiridol by administration in combination with a second agent that increases conjugative enzyme activity or glucuronyltransferase activity. This second agent should be capable of inhibiting biliary transport and may be a uridine 5'diphosphate glucuronyltransferase such as that encoded by the present sequence (designated UGT1A9). The method can be used in the treatment of cancer, gastrointestinal diseases and parasitic diseases.

Sequence 2320 BP; 608 A; 495 C; 547 G; 670 T; 0 other;

Query Match 100.0%; Score 2320; DB 24; Length 2320;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 2320; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ctcagctgactgtctgtgctgacaggggtgacagcccttctctatgtg 60
 |||||
 Db 1 ctcagctgactgtctgtgctgacaggggtgacagcccttctctatgtg 60
 |||||
 Qy 61 tgtctgtctgacactgtgcttgcagagggcaggaactactgtagtgcacatggat 120
 |||||
 Db 61 tgtctgtctgacactgtgcttgcagagggcaggaactactgtagtgcacatggat 120
 |||||
 Qy 121 gggagccactggttcacactgaggtcggtggtgagaaactcattctcaggggcatgag 180
 |||||
 Db 121 gggagccactggttcacactgaggtcggtggtgagaaactcattctcaggggcatgag 180
 |||||
 Qy 181 gtggtgtagtcacagaggtgagtgcaactggggaagatacactgaattgcacagtg 240
 |||||
 Db 181 gtggtgtagtcacagaggtgagtgcaactggggaagatacactgaattgcacagtg 240
 |||||
 Qy 241 aagacttattcaacttcataccctggaggtactgacctgacctgacctgacctgacctg 300
 |||||
 Db 241 aagacttattcaacttcataccctggaggtactgacctgacctgacctgacctgacctg 300
 |||||
 Qy 301 catgctcaatggaaagacaaagtaagaatgataatattctatttaattggttcaacaat 360
 |||||
 Db 301 catgctcaatggaaagacaaagtaagaatgataatattctatttaattggttcaacaat 360
 |||||
 Qy 361 gacattttgacttatttttcaaatgacaggagttgttttaaaagacaaaattagta 420
 |||||
 Db 361 gacattttgacttatttttcaaatgacaggagttgttttaaaagacaaaattagta 420
 |||||
 Qy 421 gaatactaaaggagagttcttttgcagtgcttctcgatccctttgataaactgtggc 480
 |||||
 Db 421 gaatactaaaggagagttcttttgcagtgcttctcgatccctttgataaactgtggc 480
 |||||
 Qy 481 ttaattgtgccaataatttctccctccctccgtggtcttgcagggggaatactttgc 540
 |||||
 Db 481 ttaattgtgccaataatttctccctccctccgtggtcttgcagggggaatactttgc 540
 |||||
 Qy 541 cactatctgaagaaggtgcacagtgccctgctctctcttctctatgtcccagaaattctc 600
 |||||
 Db 541 cactatctgaagaaggtgcacagtgccctgctctctcttctctatgtcccagaaattctc 600
 |||||
 Qy 601 ttaggtttctcagatgccaatgactttcaaggagagagtaagcaaccacatcatgacttg 660
 |||||
 Db 601 ttaggtttctcagatgccaatgactttcaaggagagagtaagcaaccacatcatgacttg 660
 |||||
 Qy 661 gaggaaacatttatgcccacgttttttcaaaaatgccctagaataagcctctgaaatt 720
 |||||
 Db 661 gaggaaacatttatgcccacgttttttcaaaaatgccctagaataagcctctgaaatt 720
 |||||
 Qy 721 ctccaaacacctgttacggagtgatgctctacagggccacacatcaattgtgttcgga 780
 |||||
 Db 721 ctccaaacacctgttacggagtgatgctctacagggccacacatcaattgtgttcgga 780
 |||||

XX DR DR XX PT PT PT XX PS XX CC CC CC CC CC CC CC CC XX SQ

WPI; 2002-075093/10.
P-PSDB: ABB04429.

Combinations of flavopiridol and an agent that increases conjugative enzyme activity or glucuronosyltransferase activity, with reduced side effects, for treating cancer -

Claim 80; Page 137-140; 145pp; English.

The present invention relates to a method of reducing the toxicity of flavopiridol by administration in combination with a second agent that increases conjugative activity in a glucuronidyltransferase activity. This second agent should be a glucuronidating pillary transport and may be a uridine 5'diphosphate glucucosyltransferase such as that encoded by the present sequence (designated UGT1A5). The method can be used in the treatment of cancer, gastrointestinal diseases and parasitic diseases.

Sequence 2320 BP; 608 A; 495 C; 547 G; 670 T; 0 other;

```
Query Match      100.0%; Score 2320; DB 24; Length 2320;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2320: Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

QY	1	ctcagctgcagttctctgtatggcttgcacaggttgacagagccctctctctatgtg	60
Db	1	ctcagctgcagttctctgtatggcttgcacaggttgacagagccctctctctatgtg	60
QY	61	tgtctgtctgacctgtggcttgcagagcgaggaaactactgttagtgcccatggat	120
Db	61	tgtctgtctgacctgtggcttgcagagcgaggaaactactgttagtgcccatggat	120
QY	121	ggagagccactggttcacacatgaagtcggtgtgtggagaaactcatctcaggggggcatgag	180
Db	121	ggagagccactggttcacacatgaagtcggtgtgtggagaaactcatctcaggggggcatgag	180
QY	181	gtggttgtgtatcgtccagaggttgagttggcaactgggaagatcactgaattgcacagtg	240
Db	181	gtggttgtgtatcgtccagaggttgagttggcaactgggaagatcactgaattgcacagtg	240
QY	241	aagacttattcaacttcataacctggagatctgcagcggaggttcaaggtctttggcc	300
Db	241	aagacttattcaacttcataacctggagatctgcagcggaggttcaaggtctttggcc	300
QY	301	catgctcaatggagaagcacaggtacgaagtatatattctctattaatgggttcatacaat	360
Db	301	catgctcaatggagaagcacaggtacgaagtatatattctctattaatgggttcatacaat	360
QY	361	gacattttgacttattttttccaaatgcagaggtttgtttaaagacaaaataatagta	420
Db	361	gacattttgacttattttttccaaatgcagaggtttgtttaaagacaaaataatagta	420
QY	421	gaatacttaagggagagttcttttatcagttttctcgatccttttgataactgtgac	480
Db	421	gaatacttaagggagagttcttttatcagttttctcgatccttttgataactgtgac	480
QY	481	ttaattgttgccaaatattctctctcctccctccgttggttttcgcagggggaatactttgc	540
Db	481	ttaattgttgccaaatattctctctcctccctccgttggttttcgcagggggaatactttgc	540
QY	541	cactattcttgaagaaggtgcacagtcgccctgctcctcttctctatgtcccagaattctc	600
Db	541	cactattcttgaagaaggtgcacagtcgccctgctcctcttctctatgtcccagaattctc	600
QY	601	ttagggttctcagatgccatgcacttccaaggagagatcaggaaccaatcatgcactgtg	660
Db	601	ttagggttctcagatgccatgcacttccaaggagagatcaggaaccaatcatgcactgtg	660
QY	661	gaggaaacatttatatgccacagttttttcaaaaatgcctcagaataatgacctcgaaatt	720
Db	661	gaggaaacatttatatgccacagttttttcaaaaatgcctcagaataatgacctcgaaatt	720

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 21, 2002, 12:45:22 ; Search time 682.45 Seconds
(without alignments)
5836.679 Million cell updates/sec

Title: US-09-835-082-3

Perfect score: 2320

Sequence: 1 ctacgtcagctctctgat.....ttctgatgtttctacaact 2320

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1736436 seqs, 858457221 residues

Total number of hits satisfying chosen parameters: 3472872

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : N_Geneseq_032802.*
1: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.*
2: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.*
3: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.*
4: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.*
5: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.*
6: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.*
7: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.*
8: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.*
9: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.*
10: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.*
11: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.*
12: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.*
13: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.*
14: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.*
15: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.*
16: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.*
17: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.*
18: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.*
19: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.*
20: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.*
21: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.*
22: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.*
23: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.*
24: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2320	100.0	2320	24	ABA05198 Human uridine 5'di
2	2320	100.0	2320	24	ABA05199 Human uridine 5'di
3	1581.4	68.2	2351	13	AAQ27369 HUG-Brl. Homo sap
4	1576.8	68.0	2368	13	AAQ27370 HUG-Brl. Homo sap
5	1551.8	66.9	1655	23	AAQ27370 HUG-Brl. Homo sap
6	1004.4	43.3	1887	24	AAQ27370 HUG-Brl. Homo sap
7	998.6	43.0	1001	21	AAH51403 DNA encoding UDP g
8	870.6	37.5	1001	21	AAH51403 Human UGT1A7 relat
9	798	34.4	1667	13	AAQ33027 Exon 5 from the UG

10	769.6	33.2	1001	21	AAH51405 Human UGT1A7 relat
11	766.6	33.0	930	21	AAZ45116 UDP-glucuronosyltr
12	741.2	31.9	951	21	AAZ45116 UDP-glucuronosyltr
13	735	31.7	735	21	AAZ45118 UDP-glucuronosyltr
14	659.8	28.4	759	21	AAZ45117 UDP-glucuronosyltr
15	578.8	24.9	874	24	ABL01468 Murine apoptosis r
16	364.2	15.7	391	20	AAZ07208 Human lung tumour
17	364.2	15.7	391	21	AAZ07208 Human lung tumour
18	364.2	15.7	391	23	AAZ07208 Human lung tumour
19	359	15.5	1001	21	AAH51406 Human UGT1A7 relat
20	353	15.2	1001	21	AAH51407 Human UGT1A7 relat
21	345.2	14.9	1854	21	AAZ95200 Human UDP-glucuron
22	345	14.9	1766	23	AAZ95200 Human UDP-glucuron
23	344.4	14.8	2107	19	AAV15900 Uridine diphospho-
24	342	14.7	380	21	AAZ98723 Human colon cancer
25	339.8	14.6	2092	21	AAZ95199 Human UDP-glucuron
26	333.2	14.4	1976	21	AAZ95206 Human UDP-glucuron
27	332.8	14.3	1650	21	AAZ95396 Human carbohydate
28	327.2	14.1	2802	23	AAZ95711 DNA encoding novel
29	325.6	14.0	2111	23	AAZ95712 DNA encoding novel
30	322.4	13.9	1636	24	AAZ24666 Human drug metabol
31	322.4	13.9	1722	22	AAZ06860 Human breast cancer
32	322.4	13.9	2974	22	AAZ46185 Human DNA encoding
33	318	13.7	1859	23	AAZ95710 DNA encoding novel
34	314.6	13.6	1713	22	AAZ06844 Human breast cancer
35	302.4	13.0	508	21	AAZ98478 Human colon cancer
36	296	12.8	1001	21	AAH51408 Human UGT1A7 relat
37	282	12.2	2025	23	AAZ91253 DNA encoding novel
38	230.4	9.9	366	21	AAA87085 Rat hepatocyte car
39	222.4	9.6	775	13	AAQ33026 Exons 3-4 from the
40	220.8	9.5	273	21	AAZ87491 Rat hepatocyte car
C 41	203	8.8	2448	23	AAZ72022 DNA encoding novel
C 42	203	8.8	2448	23	AAZ74393 DNA encoding novel
C 43	203	8.8	2448	23	AAZ78611 DNA encoding novel
C 44	203	8.8	2448	23	AAZ78773 DNA encoding novel
C 45	203	8.8	2448	23	AAZ84057 DNA encoding novel

ALIGNMENTS

RESULT 1
ABA05198
ID ABA05198 standard; cDNA; 2320 BP.
XX
AC ABA05198;
XX
DT 04-MAR-2002 (first entry)
XX
DE Human uridine 5'diphosphate glucuronyltransferase UGT1A9 cDNA #1.
XX
KW Human; uridine 5'diphosphate glucuronyltransferase enzyme; UGT1A9;
KW flavopiridol; cancer; gastrointestinal disease; parasitic infection;
KW cytosolic; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 19..1611
FT FT /*tag= a
FT FT /product= "UGT1A9"
XX
PN WO200180896-A2.
XX
PD 01-NOV-2001.
XX
PF 12-APR-2001; 2001WO-US12526.
XX
PR 21-APR-2000; 2000US-0553829.
XX
PA (ARCH-) ARCH DEV CORP.
XX
PI Retain MJ, Innocenti F, Iyer L;

68x2